THE JOURNAL
OF
CHEMICAL PHYSICS

VOLUME 9 JANUARY, 1941 NUMBER 1

Symposium on the Structure of Molecules and Aggregates of Molecules*

Introduction to the Symposium

JOHN G. KIRKWOOD
Baker Laboratory, Cornell University, Ithaca, New York
(Received October 19, 1940)

IN the previous symposia in the field of physical chemistry held by the Division of Physical and Inorganic Chemistry, molecular structure, intermolecular forces, and the kinetics of chemical reactions have been the topics of discussion. In the present symposium we return to the problem of structure. We shall be concerned in part with the structure of individual molecules and in part with the structure of interacting groups of molecules, as manifested in the liquid and solid states.

Classical physical chemistry was built upon thermodynamics. The equilibrium theory of thermodynamics was employed to correlate and interpret the physical and chemical properties of substances, without reference to their specific structure. While thermodynamics remains a frame of reference for all physico-chemical phenomena, the experimental verification of its laws is so well established, and the important task of systematizing physico-chemical data so far advanced, that its further exploitation and development now lies in the field of applied chemistry.

Modern physical chemistry is primarily concerned with problems of structure as related both to the properties of systems in equilibrium and to the mechanism of the physical and chemical processes traversed in the attainment of equilibrium. The task which it has set itself is that of relating the properties of physico-chemical systems to the structure and interactions of the molecules of which they are composed. The structure of individual molecules and the nature of intermolecular forces has been studied from the theoretical side by the methods of quantum mechanics and from the experimental side by the techniques of spectroscopy and electron diffraction. The theoretical study of the relation between molecular structure and the physico-chemical properties of systems composed of many molecules has been undertaken by the methods of statistical mechanics. The problem has been attacked from the experimental side by the techniques of x-ray crystallography, in the case of crystals, and of x-ray scattering, in the case of amorphous solids and

* The Symposium on the Structure of Molecules and Aggregates of Molecules was held, as the Fifth Annual Symposium of the Division of Physical and Inorganic Chemistry of the American Chemical Society, at Columbia University, New York, on December 30, 1940 to January 1, 1940. The arrangements for the symposium were made by a committee consisting of J. G. Kirkwood, Chairman, H. Eyring, L. S. Kassel, G. B. Kistiakowsky, L. Pauling, G. Scatchard, H. C. Urey, and J. W. Williams.

liquids. In addition much accurate experimental work has been carried out in the measurement of thermodynamic properties of particular significance to the structural theories.

The structural calculation of the thermodynamic functions of gases in the limit of zero pressure has advanced to such a stage as to be more accurate than available experimental data in many cases. In liquids and solids, the problems are much more complex, and only partial and approximate solutions can be presented for many of them at the present time. However, considerable progress has been made in the statistical mechanical study of liquids, liquid solutions, and crystalline solids, and in the interpretation of phase transitions, such as condensation, and polymorphic transformations in condensed systems.

It is the purpose of this symposium to present for discussion reports from a number of investigators, actively engaged in theoretical and experimental work on the behavior of physicochemical systems in relation to the structure and interactions of their component molecules.

JANUARY, 1941 JOURNAL OF CHEMICAL PHYSICS VOLUME 9

Molecular Distribution*

JOSEPH E. MAYER AND ELLIOTT MONTROLL**
Department of Chemistry, Columbia University, New York, New York

Various functions, by means of which the distribution of molecules in a system may be described, are defined, and the relations between them are derived. Equations, by means of which these functions may be calculated from a knowledge of the mutual potential energy between pairs of molecules, are also derived. The equations applicable to gaseous systems, which are derived by a development previously used in the theory of condensing systems, enable the calculation of the distribution functions by means of a power series in either the fugacity or the inverse volume. The coefficients of these power series are direct integrals involving the pair potentials between molecules. In condensed systems an alternative procedure is required. The equations then appear in the form of integral equations involving the unknown functions under the integral signs. These equations are not readily solved, due primarily to the difficulties of performing multiple coordinate integrations. By the use of functions involving the probabilities of population of virtual cells, the cell equation frequently used for the calculation of the thermodynamic properties of liquids is derived as a consequence of the method used here. The direct connection between the equations derived in the theory of condensing systems, and in the usual cell approximations, is thus established. The equations derived here, however, permit consistent use of the cell method to arrive at higher approximations. This development may possibly be of value in making numerical evaluations of the thermodynamic functions of liquids.

I. INTRODUCTION

FOR the sake of simplicity the discussion will be, at the outset, limited to particularly simple types of systems, although the extension to more complicated types would not involve impracticable difficulties.

Systems will be considered consisting of N identical monatomic molecules in a volume V, with volume per molecule $v=V/N$, the configurational position of the ith molecule being uniquely determined by the value of its three Cartesian coordinates, x_i, y_i, z_i, for which the symbol (i) will be used. The volume element in the three-dimensional configuration space of the ith molecule will be indicated by $d\tau_i$.

It will be assumed that the potential energy, $U\{\mathbf{N}\}$, of the system may be written as the sum of the mutual potentials of the $N(N-1)$ pairs:

$$U\{\mathbf{N}\}=\sum_{N\geq i>j\geq 1} u(\mathbf{i},\mathbf{j}), \qquad (1)$$

and that the equations of classical statistical

---

* Presented at the Symposium on the Structure of Molecules and Aggregates of Molecules at the Fifth Annual Symposium of the Division of Physical and Inorganic Chemistry of the American Chemical Society, Columbia University, New York, December 30, 1940 to January 1, 1941. Publication assisted by the Ernest Kempton Adams Fund for Physical Research of Columbia University.

** At present Sterling Fellow at Yale University.

mechanics may be used to calculate the properties of the system. It will also be assumed that the potentials $u(\mathbf{i}, \mathbf{j})$, which are functions of the distance r_{ij} only, approach zero more rapidly than r_{ij}^{-3} as r_{ij} approaches infinity, so that the integrals with which we will be concerned remain finite.

It is not necessary for what follows to assume the validity of Eq. (1). At the cost of introducing a considerably more difficult symbolism, but no essential complication, it would be possible to apply the following methods if it were only assumed that the total potential energy of the system could be developed as a sum of terms involving only pairs of molecules (as in (1)) plus terms involving three molecules each, plus terms involving four molecules each, etc., and that this development converges.

Kahn and Uhlenbeck[1] have shown that a development essentially similar to the following one may be based on quantum statistics instead of classical.

The assumption that u approaches zero more rapidly than the inverse third power of its argument, as the argument approaches infinity, is absolutely necessary in a one-component system for which all the mutual potentials have the same sign, since otherwise the internal energy of the system would be proportional to a power higher than the first of the number of molecules in the system.

It is clear that the following development could be extended to classical polyatomic molecules, whereas an extension to molecules having internal quantized states would involve only introducing summations over these states.

The authors apologize for the necessity of introducing the following complicated symbolism. The symbol $\{\mathbf{n}\}$ will be used for a set of n specified (numbered) molecules, *and* for the set of $3n$ coordinates of these molecules.

$$\{\mathbf{n}\} = x_1, y_1, z_1, x_2, \cdots, z_n,$$
$$d\tau\{\mathbf{n}\} = d\tau_1 d\tau_2 \cdots d\tau_n.$$

The molecules of the set $\{\mathbf{n}\}$ will be tacitly assumed to be numbered from 1 to n. If two sets, $\{\mathbf{n}\}$ and $\{\mathbf{m}\}$, occur in one equation, the numbering 1 to n and 1 to m will still be used, although unless specifically stated it is to be understood that the ath molecule of the set $\{\mathbf{n}\}$ is *not* the same molecule as the ath molecule of the set $\{\mathbf{m}\}$.

The symbol $\{\mathbf{v}_i\}_n$, $\nu_i \leq n$, will be used for (the coordinates of) a set of ν_i molecules, all of which are members of the set $\{\mathbf{n}\}$.

We shall have occasion to use symbols for various types of sets of such subsets $\{\mathbf{v}_i\}_n$, in which every member of the set $\{\mathbf{n}\}$ occurs in one or more of the subsets $\{\mathbf{v}_i\}_n$ of the set of subsets.

[1] B. Kahn and G. E. Uhlenbeck, Physica **5**, 399 (1938).

The symbol $\{\{\mathbf{v}_i\}_n\}_u$ will be for a set of "unconnected" subsets, $\{\mathbf{v}_i\}_n$. By unconnected we mean that each member of the set $\{\mathbf{n}\}$ occurs in one and only one subset $\{\mathbf{v}_i\}_n$ of the set of subsets. Thus in a set of unconnected subsets if a graph of n vertices is drawn with a line connecting every pair of vertices in the same subset then none of the different subsets i and j are connected by lines in the set of unconnected subsets.

The symbol $\{\{\mathbf{v}_i\}_n\}_c$ will be used for a connected set of subsets $\{\mathbf{v}_i\}_n$, by which is meant that the corresponding graph is connected in such a way that it is possible to go by a path of lines from any vertex of the set $\{\mathbf{n}\}$ to any other.

The symbol $\{\{\mathbf{v}_i\}_n\}$ will be used for an unrestricted set of subsets (in which however every member of the set $\{\mathbf{n}\}$ occurs in at least one subset $\{\mathbf{v}_i\}_n$).

Additional left-hand subscripts $\{_k\{\mathbf{v}_i\}_n\}$ will be used to indicate the number of subsets, thus $\{_k\{\mathbf{v}_i\}_n\}_u$ indicates a partition of the set $\{\mathbf{n}\}$ into k subsets, the index i which indicates the subset running from 1 to k.

II. Distribution Functions and Related Functions

The distribution functions $F_n\{\mathbf{n}\}$

The distribution function, $F_n\{\mathbf{n}\}$, a symmetrical function of the coordinates of n specified molecules, may be defined by the statement that

$$\frac{1}{V^n} F_n\{\mathbf{n}\} d\tau_1 d\tau_2 \cdots d\tau_n \qquad (2)$$

is the probability that molecule 1 is in the element of volume $d\tau_1$ at the position x_1, y_1, z_1, simultaneously molecule 2 in the element $d\tau_2$ at x_2, y_2, z_2, $\cdots$ and molecule n in the element $d\tau_n$ at x_n, y_n, z_n.

The functions so defined are normalized so that their average values are unity:

$$\frac{1}{V^n} \int_V \int \cdots \int F_n\{\mathbf{n}\} d\tau\{\mathbf{n}\} = 1. \qquad (3)$$

Owing to their definitions, (2), as probabilities, the distribution functions for various values of n are related by the equations:

$$F_m\{\mathbf{m}\} = \frac{1}{V^{n-m}} \int_V \int \cdots \int F_n\{\mathbf{n}\} d\tau_{m+1} \cdots d\tau_n, \qquad (m < n). \quad (4)$$

The function $F_1(1)$ is a constant, equal to unity,

$$F_1(1)=1. \tag{5}$$

All the functions are functions only of the distances between the molecules, so that they are invariant under the three-dimensional translation and rotation group.

In one-phase systems consisting of liquids, gases, or glasses, for which long range order is absent, the distribution functions approach unity, to within terms of order $1/N$ when all the distances between all of the molecules of the set $\{\mathbf{n}\}$ become large. If the coordinates of the n molecules have such values that the distances between all molecules belonging to different subsets of the set of unconnected subsets $\{_k\{\mathbf{v}_i\}_n\}_u$ are large then

$$F_n\{_k\{\mathbf{v}_i\}_n\}_u=\prod_{i=1}^{i=k} F_{\nu_i}\{\mathbf{v}_i\}_n, \quad (r_{ij}\gg v^{\frac{1}{3}}). \tag{6}$$

The auxiliary functions, $g_\nu\{\mathbf{v}\}$

Symmetrical functions $g_\nu\{\mathbf{v}\}$ of the coordinates of ν molecules may be defined in terms of the distribution functions $F_n\{\mathbf{n}\}$ by either of the two reciprocal sets of equations:

$$F_n\{\mathbf{n}\}=\sum_{\{_k\{\mathbf{v}_i\}_n\}_u}\prod_{i=1}^{i=k} g_{\nu_i}\{\mathbf{v}_i\}_n \tag{7}$$

or

$$g_\nu\{\mathbf{v}\}=\sum_{\{_k\{\mathbf{n}_i\}_\nu\}_u}(-)^{k-1}(k-1)!\prod_{i=1}^{i=k} F_{n_i}\{\mathbf{n}_i\}_\nu, \tag{8}$$

in which the summation in each case goes over all unconnected sets of subsets, and the product over all k subsets. (In this and some subsequent equations the notation $\sum_{\{_k\{\mathbf{v}_i\}_n\}_u}$ is used to indicate summation over all possible sets of subsets restricted to the type indicated. The symbol $\{_k\{\mathbf{v}_i\}_n\}_u$ is not set inferior to the $\sum$ symbol merely to avoid typesetting difficulties.)

Since, from (7) or (8), $g_1(1)=F_1(1)$, it follows from (5) that

$$g_1(1)=1. \tag{9}$$

In systems containing one noncrystalline phase, lacking long range order of any sort, a comparison of Eqs. (6) with (7) shows that $g_\nu\{\mathbf{v}\}$ approaches zero if all of the $\nu-1$ distances of any *one* of the molecules to all of the others of the set becomes very large.

It can readily be shown, by a comparison of (4) with (8) that

$$\int_V g_\nu\{\mathbf{v}\}d\tau_i=0, \tag{10}$$

where the integration is carried out over the coordinates of any one molecule. In view of the fact that integrand of (10) approaches zero in value when the coordinates of the ith particle are far distant from those of the other $\nu-1$ particles it is tempting to assume that the only contribution to the integral comes from that part of the configuration space where all the particles are close together, and to write the limits infinity in Eq. (10). Unfortunately this is not correct.

The situation is most clearly visualized intuitively in the case of a system composed of two hard spheres of radius r_0 each in a volume V. In this case $F_2(r_{12})$ is zero if $r_{12}\leqq 2r_0$, and $V/[V-(32/3)r_0^3]\leqq 1+(32r_0^3/3V)$ everywhere else. The function $g_2(r_{12})=F_2(r_{12})-1$ is then -1 for $r_{12}\leqq 2r_0$ and $32r_0^3/3V$ elsewhere. Equation (10) is then satisfied, provided the integration is extended over the volume V of the system. If V becomes large, r_0^3/V approaches zero, and $g_2(r_{12})$ approaches zero for large values of its argument, as is generally the case in systems of large V.

Of course the actual distribution function for N hard spheres in the volume V will be different from that of the above case, but this example suffices to demonstrate the error of assuming that $g_2(r_{12})$ approaches zero for large values of its argument sufficiently rapidly to substitute infinite limits in the integral of Eq. (10).

The auxiliary functions $\mathbf{h}_\nu\{\mathbf{v}\}$

Another convenient set of symmetrical functions of the coordinates of ν particles are the $h_\nu\{\mathbf{v}\}$, defined as

$$\ln F_n\{\mathbf{n}\}=\sum_{\{\mathbf{v}\}_n} h_\nu\{\mathbf{v}\}_n, \tag{11}$$

$$h_\nu\{\mathbf{v}\}=\sum_{\{\mathbf{n}\}_\nu}(-)^{\nu-n}\ln F_n\{\mathbf{n}\}_\nu, \tag{12}$$

with $\ln F_1(1)=0=h_1(1)$.

These functions, $h_\nu\{\mathbf{v}\}$, have, in common with the $g_\nu\{\mathbf{v}\}$, the property of becoming zero (in one-phase glass, liquid or gaseous systems) if the distances between any one of the set and the other $\nu-1$ particles become large.

The function $h_3(1,2,3)$ determines the deviation of $F_3(1,2,3)$ from the simple product $F_2(2,3)F_2(1,3)F_2(1,2)$ which it may be expected to approach in value. The higher members $h_\nu\{\mathbf{v}\}$ have similar interpretations.

The $h_\nu\{\mathbf{v}\}$ and $g_\nu\{\mathbf{v}\}$ are connected by the equation:

$$g_\nu\{\mathbf{v}\} = \sum_{(\{\mathbf{v}_i\}_\nu)_c} \prod_i [(\exp h_{\mu_i}\{\mathbf{v}_i\}_\nu) - 1]$$

$$= \sum_{(\{\mathbf{v}_i\}_\nu)} \prod_i \sum_{n_i \geq 1} \frac{[h_{\mu_i}\{\mathbf{v}_i\}_\nu]^{n_i}}{n_i!}, \qquad (13)$$

in which the summation runs over all connected sets of the subsets $\{\mathbf{v}_i\}_\nu$, and the products over all subsets i.

Distribution functions in two-phase systems

The two phases will be referred to as a and b, and the phase a will be assumed to be the denser. The symbols

$$\rho_a = N_a/V_a, \quad \rho_b = N_b/V_b, \quad \rho = N/V$$

with

$$\rho_a > \rho > \rho_b,$$

and

$$N_a + N_b = N, \quad V_a + V_b = V,$$

will be used for the densities, numbers of molecules, and volumes of the two phases and of the total system. The mass fractions of the phases are related to the densities by

$$\frac{N_a}{N} = \frac{\rho_a}{\rho}\frac{\rho - \rho_b}{\rho_a - \rho_b}, \quad \frac{N_b}{N} = \frac{\rho_b}{\rho}\frac{\rho_a - \rho}{\rho_a - \rho_b}, \qquad (14)$$

and the volume fractions by

$$\frac{V_a}{V} = \frac{\rho - \rho_b}{\rho_a - \rho_b}, \quad \frac{V_b}{V} = \frac{\rho_a - \rho}{\rho_a - \rho_b}. \qquad (15)$$

We consider a two-phase system in which the regions occupied by each of the phases are contiguous, and moderately compact in shape. Going to the limit of a system of infinite size, the members of a finite set of molecules which are at finite distances from each other will be either *all* in phase a or *all* in phase b. The distribution functions, F_n, for such a system (assuming that all regions have equal probability of being occupied by any one of the phases, which will only be the case in the absence of gravitational fields) will be a linear combination of the two distribution functions, ${}_aF_n$ and ${}_bF_n$ of the two phases, in which, however, the sum of the coefficients is not necessarily unity.*

The correct linear combination may be seen rather readily to be

$$F_n\{\mathbf{n}\} = \frac{\rho - \rho_b}{\rho_a - \rho_b}\left(\frac{\rho_a}{\rho}\right)^n {}_aF_n\{\mathbf{n}\} + \frac{\rho_a - \rho}{\rho_a - \rho_b}\left(\frac{\rho_b}{\rho}\right)^n {}_bF_n\{\mathbf{n}\}. \qquad (16)$$

Although Eq. (5) that $F_1 = 1$, is obeyed, $F_n\{\mathbf{n}\}$ does not, in general, approach unity in value as the distances between all of the molecules become large, in this two-phase system. A type of long range order is necessarily present in the two-phase system even if both phases individually lack it.

Equation (6) does not apply in the case of the two-phase system, nor do the functions $g_\nu\{\mathbf{v}\}$ and $h_\nu\{\mathbf{v}\}$ defined by Eqs. (7) and (11) approach zero as the distance between one of the molecules and all of the $\nu - 1$ others becomes large. Thus although the distribution functions for liquid and gas are similar in all the characteristics discussed so far, those of the intermediate two-phase region are qualitatively entirely different from those of either of the two pure phases.

Population probability functions[2] $P(\{\mathbf{R}\}\{\nu\})$

The molecular distribution functions F_n determine the probability densities for *specified* molecules to occupy certain positions in space. For some purposes it is more convenient to operate with functions determining the probabilities that certain fixed regions of the three-dimensional volume space shall be occupied simultaneously by specified numbers of molecules.

Functions $P(\{\mathbf{R}\}\{\nu\})$ may be defined by the statement that

$$P(\{\mathbf{R}\}\{\nu\}) = P(\mathbf{1}, \cdots, \mathbf{r}, \cdots, \mathbf{R}, \nu_1, \cdots, \nu_r, \cdots, \nu_R) \qquad (17)$$

is the probability that in the R non-overlapping regions, $\mathbf{1}, \cdots, \mathbf{R}$, there shall be exactly ν_1 molecules in the first region, $\mathbf{1}$, and simultaneously . . . ν_R molecules in the last region, $\mathbf{R}$.

* The authors are indebted to Professor Norbert Wiener for first pointing out this fact to them.

[2] The relations between these functions and the distribution functions has been discussed by Professor Norbert Wiener in a manuscript which it has been the privilege of the authors to see. The derivations given here differ slightly from those of Wiener, but are essentially inspired by his.

It is clear from the definition (17) that

$$P(1, \cdots, \mathbf{R}, \nu_1, \cdots, \nu_R) = \sum_{\nu_s \geq 0} P(1, \cdots, \mathbf{R}, \mathbf{s}, \nu_1, \cdots, \nu_R, \nu_s), \tag{18}$$

and that

$$P(1, \cdots, \mathbf{R}-1, \mathbf{R}+\mathbf{s}, \nu_1, \cdots, \nu_{R-1}, \nu_{R+s}) = \sum_{\substack{\nu_R, \nu_s \\ \nu_R+\nu_s=\nu_{R+s}}} P(1, \cdots, \mathbf{R}-1, \mathbf{R}, \mathbf{s}, \nu_1, \cdots, \nu_{R-1}, \nu_R, \nu_s) \cdot \tag{19}$$

in which the region $\mathbf{R}+\mathbf{s}$ appearing in the function on the left consists of the sum of the two portions of space occupied by the two regions $\mathbf{R}$ and $\mathbf{s}$ in the function on the right.

These functions P are related to certain integrals of the F_n's over the set of regions $\{\mathbf{R}\}$. The integral

$$\phi'(1, \cdots, \mathbf{R}, \nu_1, \cdots \nu_R) = \frac{1}{V^n} \int\int \cdots \int_{\{\mathbf{R}\}\{\nu\}} F_n\{\mathbf{n}\} d\tau\{\mathbf{n}\}, \tag{20}$$

with $\nu_r \geq 1$ and $\sum_r \nu_r = n$, in which the integration of the first ν_1 molecules is extended over the region **1**, of the ν_1+1th to the $\nu_1+\nu_2$th molecule over **2**, etc., gives the probability that the specified molecules numbered 1 to ν_1 are in the region **1**, simultaneously the molecules ν_1+1 to $\nu_1+\nu_2$ inclusive in the region **2**, $\cdots$, and molecules $n-\nu_R+1$ to n inclusive in the region $\mathbf{R}$.

The total number of ways that one can choose, out of N numbered molecules R sets of molecules with ν_r in the rth set is $N!/(N-n)!\prod \nu_r!$. For $n \ll N$ one may write $N!/(N-n)! = N^n$ and $N^n/V^n = 1/v^n$. The functions

$$\begin{aligned} \phi(\{\mathbf{R}\}\{\nu\}) &= \frac{N!}{(N-n)!\pi\nu_r!} \phi'(\{\mathbf{R}\}\{\nu\}) \\ &= \{\prod_{r=1}^{r=R} \nu_r!\}^{-1} \frac{1}{v^n} \int\int \cdots \int_{\{\mathbf{R}\}\{\nu\}} F_n\{\mathbf{n}\} d\tau\{\mathbf{n}\}, \quad (\sum_{r=1}^{r=R} \nu_r = n), \end{aligned} \tag{21}$$

represent the sum over all possible choices of the numbered molecules of the probabilities that in the set of regions $\{\mathbf{R}\}$ there be ν_r specified molecules in the rth region $\mathbf{r}$.

If there are exactly μ_r molecules in the region $\mathbf{r}$, $\mu_r \geq \nu_r$, there will be $\mu_r!/(\mu_r-\nu_r)!\nu_r!$ different sets of ν_r numbered molecules in the region. One may consequently write

$$\phi(\{\mathbf{R}\}\{\nu\}) = \sum_{\mu_r \geq \nu_r \{\mu\}} \left[\prod_{r=1}^{r=R} \frac{\mu_r!}{(\mu_r-\nu_r)!\nu_r!} \right] P(\{\mathbf{R}\}\{\mu\}). \tag{22}$$

No provision has been made for defining the functions ϕ when one of the ν_r's is equal to zero. In view of Eqs. (18) and (22) it is logical to define

$$\phi(1, \cdots, \mathbf{R}, \mathbf{s}, \nu_1, \cdots, \nu_R, 0) = \phi(1, \cdots, \mathbf{R}, \nu_1, \cdots, \nu_R), \tag{23}$$

and

$$\phi(1, \cdots \mathbf{R}, 0, \cdots, 0) = 1. \tag{24}$$

With these definitions the inverse of the set of equations, (22), is given by the set

$$P(\{\mathbf{R}\}\{\mu\}) = \sum_{\nu_r \geq \mu_r \{\nu\}} \left[\prod_{r=1}^{r=R} \frac{\nu_r!(-)^{\nu_r-\mu_r}}{(\nu_r-\mu_r)!\mu_r!} \right] \phi(\{\mathbf{R}\}\{\nu\}). \tag{25}$$

If functions $\gamma(\{\mathbf{R}\}\{\kappa\})$ are defined as integrals over the $g_\nu\{\nu\}$ similarly to the definition (21) of the ϕ's in terms of the F_n's:

$$\gamma(\{\mathbf{R}\}\{\kappa\}) = \{\prod_{r=1}^{r=R} \kappa_r!\}^{-1} \frac{1}{v^\nu} \int\int \cdots \int_{\{\mathbf{R}\}\{\kappa\}} g_\nu\{\nu\} d\tau\{\nu\}, \quad (\sum_{r=1}^{r=R} \kappa_r = \nu), \tag{26}$$

with the relationships equivalent to (23) and (24), then

$$P(\{\mathbf{R}\}\{\mathfrak{y}\}) = \{\prod_{r=1}^{r=R} \mu_r!\}^{-1} \lim_{x_r \to -1} \{[\prod_{r=1}^{r=R} (\partial/\partial x_r)^{\mu_r}] \exp \sum_{k_r \geq 0} \gamma(\{R\}\{\kappa\}) \prod_{r=1}^{r=R} x_r^{\kappa_r}\}. \tag{27}$$

III. The General Physical Equations

In this section the general physical equations determining the distribution functions in terms of the mutual potentials of the pairs of molecules will be derived.

The configuration integral

$$\mathbf{Q}_r(N, V, T) = \int_V \int \cdots \int \exp[-U\{\mathbf{N}\}/kT] d\tau\{\mathbf{N}\}, \tag{28}$$

is related to the Helmholtz free energy, A, by the equation

$$A = -kT \ln \left[\left(\frac{2\pi mkT}{h^2} \right)^{3N/2} \frac{\mathbf{Q}_r(N, V, T)}{N!} \right]. \tag{29}$$

The chemical potential, μ, is

$$\mu = \left(\frac{\partial A}{\partial N} \right)_{V,T} = -kT \ln \left[\left(\frac{2\pi mkT}{h^2} \right)^{\frac{3}{2}} \frac{\mathbf{Q}_r(N+1, V, T)}{(N+1)\mathbf{Q}_r(N, V, T)} \right]. \tag{30}$$

The fugacity, Z, in molecules per unit volume, may be defined by the equation

$$\mu = -kT \ln \left(\frac{2\pi mkT}{h^2} \right)^{\frac{3}{2}} \frac{1}{Z}, \tag{31}$$

so that by comparison with Eq. (30) it is seen that

$$Z = \frac{(N+1)\mathbf{Q}_r(N, V, T)}{\mathbf{Q}_r(N+1, V, T)}. \tag{32}$$

The distribution function, $F_N\{\mathbf{N}\}$, of *all* of the molecules of the system is given by the equation

$$F_N\{\mathbf{N}\} = V^N \frac{1}{\mathbf{Q}_r(N, V, T)} \exp[-U\{\mathbf{N}\}/kT], \tag{33}$$

and from (4) that for $F_\lambda\{\mathfrak{a}\}$ by

$$F_\lambda\{\mathfrak{a}\} = V^\lambda \frac{1}{\mathbf{Q}_r(N, V, T)} \int_V \int \cdots \int \exp[-U\{\mathbf{N}\}/kT] d\tau_{\lambda+1} \cdots d\tau_N. \tag{34}$$

The exponent of $-U\{\mathfrak{a}\}/kT$, which occurs as a factor of the integrand, may be taken out from under the integral sign, as the integration extends only over the molecules other than those of the set $\{\mathfrak{a}\}$. With the assumption of Eq. (1) the remaining term of the integrand is a product of the exponent of terms involving only pairs of molecules. The running index κ will be used for molecules of the set $\{\mathfrak{a}\}$, $\lambda \geq \kappa \geq 1$, and the indices i, j and k for molecules of the set $\{\mathbf{N}-\mathfrak{a}\}$, $N \geq i > j \geq \lambda+1$, $N \geq k \geq \lambda+1$. By use of the functions

$$f_{ij} = f(\mathbf{i}, \mathbf{j}) = \exp[-u(\mathbf{i}, \mathbf{j})/kT] - 1, \tag{35}$$

the remaining part of the integrand may be written either as

$$\prod_{\substack{N\geq i>j\geq\lambda+1\\ N\geq k\geq\lambda+1\\ \lambda\geq\kappa\geq 1}}(1+f_{ij})(1+f_{\kappa k})=\sum_{\substack{\nu_{ij}=0,1\\ \nu_{\kappa k}=0,1}}\prod_{\substack{N\geq i>j\geq\lambda+1\\ N\geq k\geq\lambda+1\\ \lambda\geq\kappa\geq 1}}f_{ij}^{\nu_{ij}}f_{\kappa k}^{\nu_{\kappa k}},$$

or as

$$\Big[\prod_{\substack{N\geq i\geq\lambda+1\\ \lambda\geq\kappa\geq 1}}(1+f_{\kappa i})\Big]\big[\exp(-U\{\mathbf{N}-\boldsymbol{\lambda}\}/kT)\big]=\sum_{\nu_{\kappa i}=0,1}\Big[\prod_{\substack{N\geq i\geq\lambda+1\\ \lambda\geq\kappa\geq 1}}f_{\kappa i}^{\nu_{\kappa i}}\Big]\big[\exp(-U\{\mathbf{N}-\boldsymbol{\lambda}\}/kT)\big].$$

The expression for the distribution functions $F_\lambda\{\boldsymbol{\lambda}\}$ may then be written either as

$$F_\lambda\{\boldsymbol{\lambda}\}=V\big[\exp(-U\{\boldsymbol{\lambda}\}/kT)\big]\frac{1}{\mathbf{Q}_\tau(N,V,T)}\int_V\int\cdots\int\sum_{\substack{\nu_{ij}=0,1\\ \nu_{\kappa k}=0,1}}\prod_{\substack{N\geq i>j\geq\lambda+1\\ N\geq k\geq\lambda+1\\ \lambda\geq\kappa\geq 1}}f_{ij}^{\nu_{ij}}f_{\kappa k}^{\nu_{\kappa k}}d\tau\{\mathbf{N}-\boldsymbol{\lambda}\},\quad(36)$$

or as

$$F_\lambda\{\boldsymbol{\lambda}\}=V\big[\exp(-U\{\boldsymbol{\lambda}\}/kT)\big]\frac{1}{\mathbf{Q}_\tau(N,V,T)}\int_V\int\cdots\int\big[\exp-U\{\mathbf{N}-\boldsymbol{\lambda}\}/kT\big]\times\sum_{\nu_{\kappa i}=0,1}\prod_{\substack{N\geq i\geq\lambda+1\\ \lambda\geq\kappa\geq 1}}f_{\kappa i}^{\nu_{\kappa i}}d\tau\{\mathbf{N}-\boldsymbol{\lambda}\}.\quad(36')$$

IV. Distribution Functions for Gases

Development as a power series in Z

The further use of the form (36) for evaluation of the distribution functions $F_\lambda\{\boldsymbol{\lambda}\}$ appears to lead to series which diverge if the system exists partially or wholly as a condensed phase, so that the subsequent equations of this section are applicable, presumably, only to gases.*

The integrand of (36) is a sum of products. In each product the molecules may be assigned to "clusters," consisting of one or more molecules, for which the corresponding graphs[3] are singly or more connected. Consider one special product of the sum in which the clusters containing molecules of the set $\{\boldsymbol{\lambda}\}$ contain also L specified molecules of the set $\{\mathbf{N}-\boldsymbol{\lambda}\}$. The integration over the coordinates of the remaining $N-\lambda-L$ *not* connected in the same cluster to any of the set $\{\boldsymbol{\lambda}\}$ leads to a constant factor in the contribution of this term to the integral. Now consider the sum of *all* products in which the same L molecules are connected *in the same manner* to those of the set $\{\boldsymbol{\lambda}\}$, but in which all possible products of f's of the remaining $N-\lambda-L$ molecules occur. The contribution of the integral over the coordinates of these $N-\lambda-L$ molecules in the sum of these terms is just the factor $\mathbf{Q}_\tau(N-\lambda-L,V,T)$.

In the product originally chosen there are, let us say, θ clusters containing molecules of the set $\{\boldsymbol{\lambda}\}$. The running index α, $\theta\geq\alpha\geq 1$, will be used to indicate the different clusters. In this product, then, the subset $\{\boldsymbol{\nu}_\alpha\}_\lambda$ is in the same cluster with the subset $\{\boldsymbol{l}_\alpha\}_L$ of the set $\{\mathbf{L}\}$. We now sum over all products in which the same subsets $\{\boldsymbol{\nu}_\alpha\}_\lambda$ are in one cluster with the same subsets $\{\boldsymbol{l}_\alpha\}_L$ for all values of α, $\sum_\alpha\nu_\alpha=\lambda$, $l_\alpha\geq 0$, $\sum_\alpha l_\alpha=L$.

To do this functions of the coordinates of ν particles, $B_{\nu l}\{\boldsymbol{\nu}\}$, are defined (for $\nu\geq 1$, $l\geq 1$) as

$$B_{\nu l}\{\boldsymbol{\nu}\}=\frac{1}{l!}\int\int\cdots\int\sum\prod_{\substack{l\geq i>j\geq 1\\ l\geq k\geq 1\\ \nu\geq\kappa\geq 1}}f_{ij}f_{\kappa k}d\tau\{\boldsymbol{l}\}\quad(37)$$

sum over all connected products

The special cases for $l=0$,

$$B_{1,0}(1)=1,\quad B_{\nu,0}\{\boldsymbol{\nu}\}=0\text{ if }\nu>1,\quad(37')$$

are to be included in the above set of functions.

The integral over the sum of all the products limited by having the same associated subsets $\{\boldsymbol{\nu}_\alpha\}_\lambda$ and $\{\boldsymbol{l}_\alpha\}_L$ in the cluster α, is just

$$\mathbf{Q}_\tau(N-\lambda-L,V,T)\prod_{\theta\geq\alpha\geq 1}(l_\alpha!B_{\nu_\alpha l_\alpha}\{\boldsymbol{\nu}_\alpha\}_\lambda).$$

* Professor Norbert Wiener has suggested that whereas equations of the type developed here are probably invalid in the intermediate range of condensation, they may again become asymptotically valid at the higher densities of the liquid phase.

[3] The method used here follows closely that previously employed by Mayer and S. F. Harrison, J. Chem. Phys. 6, 87 (1938).

There are $(N-\lambda)!/(N-\lambda-L)!\prod l_\alpha!$ ways of assigning the numbered $N-\lambda$ molecules of the set $\{\mathbf{N}-\boldsymbol{\lambda}\}$ to the subsets $\{\boldsymbol{l}_\alpha\}_L$, so there will be this many terms of the above type occurring in the integral (36).

For $\lambda \ll N$, $L \ll N$, we write

$$(N-\lambda)!/(N-\lambda-L)! = N^L,$$

and

$$N^L \mathbf{Q}_\tau(N-\lambda-L,\ V,\ T)/\mathbf{Q}_\tau(N,\ V,\ T) = Z^{L+\lambda}/N^\lambda,$$

and

$$V^\lambda Z^{L+\lambda}/N^\lambda = (vZ)^\lambda Z^L,$$

from Eq. (32).

Equation (36) may then be written as

$$F_\lambda\{\boldsymbol{\lambda}\} = [\exp(-U\{\boldsymbol{\lambda}\}/kT)] \sum_{\{\{\boldsymbol{\nu}_\alpha\}\lambda\}_u} \times \prod_\alpha [(vZ)^{\nu_\alpha} \sum_{l\geq 0} B_{\nu_\alpha l}\{\boldsymbol{\nu}_\alpha\}_\lambda Z^l]. \quad (38)$$

By defining

$$G_\nu\{\boldsymbol{\nu}\} = (vZ)^\nu \sum_{l\geq 1} B_{\nu l}\{\boldsymbol{\nu}\} Z^l, \quad \nu > 1, \quad (39)$$

and

$$G_1(1) = \sum_{l\geq 0} v B_{1l} Z^{l+1}, \quad (39')$$

which latter is a constant, independent of the value of the coordinates of the molecule, we obtain

$$F_\lambda\{\boldsymbol{\lambda}\} = [\exp(-U\{\boldsymbol{\lambda}\}/kT)] \sum_{\{\{\boldsymbol{\nu}_\alpha\}\lambda\}_u} \times \prod_\alpha G_{\nu_\alpha}\{\boldsymbol{\nu}_\alpha\}_\lambda. \quad (40)$$

For the special case that $\lambda = 1$, from Eq. (5), it follows that

$$F_1(1) = 1 = G_1(1). \quad (41)$$

Equations (37) and (37′) for $\nu = 1$ show that $B_{1,\,l-1}$ is just l times the cluster integral b_l previously[3] introduced in the theory of the imperfect gas, and defined by the equation

$$b_l = \frac{1}{V}\frac{1}{l!}\int\int\cdots\int \sum_{\text{sum over all connected products}} \prod_{l\geq i>j\geq 1} f_{ij}\, d\tau\{\boldsymbol{l}\} \quad (42)$$

so that (39′) and (41) become

$$\sum_{l\geq 1} v l b_l Z^l = 1, \quad (43)$$

the fundamental equation for the determination of the fugacity Z in the imperfect gas theory.

The functions $G_\nu\{\boldsymbol{\nu}\}$ are seen to have the same relation, Eq. (40), to $F_\lambda\{\boldsymbol{\lambda}\}\exp(U\{\boldsymbol{\lambda}\}/kT)$, that the functions $g_\nu\{\boldsymbol{\nu}\}$ do to $F_\lambda\{\boldsymbol{\lambda}\}$, Eq. (7).

The development, at least as presented here, is valid only if the series (39) converge, which will be the case as long as the volume is large enough for the system to remain gaseous. The integrals $B_{\nu l}\{\boldsymbol{\nu}\}$, for finite values of l, approach zero if any of the molecules of the set $\{\boldsymbol{\nu}\}$ is far distant from all the others of the set, or if the set divides naturally into two or more subsets, the molecules of which are all far distant from those in other subsets. The functions $G_\nu\{\boldsymbol{\nu}\}$ have, then, also this property if the series (39) converge. The distribution functions, $F_\lambda\{\boldsymbol{\lambda}\}$ are therefore seen to obey Eq. (6).

Since the factor $\exp(-U\{\boldsymbol{\lambda}\}/kT)$ becomes zero if any two molecules approach each other too closely, it is seen that the distribution functions do so also. At extremely large volumes of the system, for which $Z \cong 1/v$ becomes small, the functions $G_\nu\{\boldsymbol{\nu}\}$, for $\nu > 1$, approach zero in value, and the functions $F_\lambda\{\boldsymbol{\lambda}\}$ approach $\exp(-U\{\boldsymbol{\lambda}\}/kT)$ in value.

Development as an inverse power series in v

The integrand of the functions $B_{\nu l}\{\boldsymbol{\nu}\}$, Eq. (37), consists of a sum of products, each of which may, in the usual manner, be represented by a graph. All singly or more connected graphs of $l+\nu$ vertices, with no more than one line between two given vertices, are represented in the graphs describing the terms of this integrand.

Among the graphs representing these products will be many in which one or more of the vertices of the set $\{\boldsymbol{l}\}$ are but singly connected to the vertices of the rest of the graph, and the integration over the coordinates of these molecules will contribute only a constant factor in the integral of the product. It is therefore possible to "reduce" the integrals $B_{\nu l}\{\boldsymbol{\nu}\}$, and represent each of them as a sum of products of simpler integrals. This reduction of the integrals is essentially similar to the representation of the cluster integrals b_l as a sum of products of the "irreducible integrals" β_k in the theory of the imperfect gas,[3] so that it will be only necessary to consider here the case that $\nu > 1$.

We define new functions $H_m\{\boldsymbol{\nu}\}$, for $\nu > 1$, similarly to the $B_{\nu l}\{\boldsymbol{\nu}\}$ of Eq. (37), but with a

greater restriction on the type of products occurring in the integrand, by the equation

$$H_{\nu m}\{\nu\}=\frac{1}{m!}\int\int\cdots\int\sum\prod_{\substack{m\geq i>j\geq 1\\ m\geq k\geq 1\\ \nu\geq\kappa\geq 1}} f_{ij}f_{\kappa k}\,d\tau\{\mathbf{m}\},\qquad(\nu>1).\quad(44)$$

sum over all connected products for which each molecule of the set $\{\mathbf{m}\}$ is connected by an independent path to at least two of the set $\{\nu\}$.

The irreducible integrals β_k, used previously,

$$\beta_k=\frac{1}{V}\frac{1}{k!}\int\int\cdots\int\sum_{\substack{\text{sum over all doubly or more}\\ \text{connected products}}}\prod_{k+1\geq i>j\geq 1} f_{ij}\,d\tau\{\mathbf{k}+1\}\quad(45)$$

will also be introduced.

The sum of certain of the products in the integrand of (37), characterized by having the same sets of k molecules each, with each set but singly connected in the same way to the rest of the graph, and the same set $\{\mathbf{m}\}_l$ each, connected to at least two of the set $\{\nu\}$, will lead to a product of the type

$$[\prod_k (k!\beta_k)^{n_k}]m!H_{\nu m}\{\nu\},\quad\text{with}\quad\sum_k kn_k+m=l$$

after integration. By a nasty, but straightforward, combinatorial argument it can be shown that the number of sets of products of the above type leading to this integral is

$$[(m+\nu)l!/(l+\nu)m!]\prod_k[(l+\nu)^{n_k}/(k!)^{n_k}n_k!].$$

It follows that one may write

$$B_{\nu l}\{\nu\}=\sum_{m\geq 1}H_{\nu m}\{\nu\}\Big[\sum_{\sum kn_k=l-m}\prod_k\frac{m+\nu}{l+\nu}\frac{[(l+\nu)\beta_k]^{n_k}}{n_k!}\Big],\qquad(\nu>1).\quad(46)$$

Since it has been found previously[3] that

$$b_{l+\nu}=\frac{1}{(l+\nu)^2}\sum_{\sum_k kn_k=l+\nu-1}\prod_k\frac{[(l+\nu)\beta_k]^{n_k}}{n_k!},\quad(47)$$

one may write (46) as

$$B_{\nu l}\{\nu\}=\sum_{m\geq 1}(m+\nu)\frac{\partial b_{l+\nu}}{\partial\beta_{m+\nu-1}}H_{\nu m}\{\nu\},$$

and from (39)

$$G_\nu\{\nu\}=(vZ)^\nu\sum_{l\geq 1}B_{\nu l}\{\nu\}Z^l$$

$$=v^\nu\sum_{m\geq 1}(m+\nu)\Big[\sum_{l\geq\nu+1}Z^l\frac{\partial b_l}{\partial\beta_{m+\nu-1}}H_{\nu m}\{\nu\}\Big].\quad(48)$$

The equations

$$\sum_{l\geq 1}vb_lZ^l=1-\sum_{k\geq 1}\frac{k}{k+1}\beta_kv^{-k},\quad(49)$$

and

$$\ln Z=-\ln v-\sum_{k\geq 1}\beta_kv^{-k},\quad(50)$$

both of which have been derived in making the developments for the imperfect gas,[3] may be used to derive that

$$\sum_{l\geq 1}Z^l\frac{\partial b_l}{\partial\beta_k}=-\frac{1}{v}\frac{\partial}{\partial\beta_k}\sum_{l\geq 1}vb_lZ^l-\frac{1}{v}\sum_{l\geq 1}ivb_lZ^l\frac{\partial\ln Z}{\partial\beta_k}$$

$$=\frac{1}{k+1}v^{-(k+1)}.\quad(51)$$

Since b_l for $l<k+1$ is independent of β_k, this equation also holds for the sum occurring in Eq. (48). One finds, then,

$$G_\nu\{\nu\}=\sum_{m\geq 1}v^{-m}H_{\nu m}\{\nu\},\quad\nu>1;\quad G_1(1)=1.\quad(52)$$

Equations for the auxiliary functions $h_\kappa\{\nu\}$

Equation (52) may be reduced still further. The functions $H_{\nu m}\{\nu\}$ may also be represented as a sum of products. If, in the graph corresponding to a given product of the integrand of (44) for $H_{\nu m}\{\nu\}$, the set of vertices $\{\mathbf{m}\}$ can be partitioned into θ subsets $\{\mathbf{n}_i\}_m$, $\theta\geq i\geq 1$, such that vertices of the same subset are connected by paths not passing through vertices of the set $\{\nu\}$, whereas vertices of different subsets, $\{\mathbf{n}_i\}_m$ and $\{\mathbf{n}_j\}_m$, are connected *only* through vertices of the set $\{\nu\}$, then the integral of the sum of all products thus limited may be written as

$$\prod_{i=1}^{i=\theta}(n_i!h_{\mu_i n_i}\{\mathbf{\mu}_i\}_\nu)^{k_i},\quad\sum_{i=1}^{i=\theta}n_ik_i=m,$$

in which the θ subsets $\{\mathfrak{y}_i\}_\nu$ form a connected set $\{_\theta\{\mathfrak{y}_i\}_\nu\}_c$, and may, indeed, include two identical subsets, $\{\mathfrak{y}_j\} \equiv \{\mathfrak{y}_i\}$ if $n_j \neq n_i$. The functions $h_{\mu n}\{\mathfrak{y}\}$ are defined in a manner that differs from the definition of $B_{\nu l}$ or $H_{\nu m}$ only in that the products included in the integrand are still more restricted than in either of the latter two functions:

$$h_{\mu n}\{\mathfrak{y}\} = \frac{1}{n!}\int\int\cdots\int \sum \prod_{\substack{n \geq i > j \geq 1 \\ n \geq k \geq 1 \\ \mu \geq \kappa \geq 1}} f_{ij} f_{\lambda k} d\tau\{\mathfrak{n}\} \tag{53}$$

sum over all connected products in which the vertices of the set $\{\mathfrak{n}\}$ are connected independently of these of the set $\{\mathfrak{y}\}$, and each vertex of the set $\{\mathfrak{n}\}$ is connected by independent paths to at least 2 vertices of the set $\{\mathfrak{y}\}$.

There are clearly just $m!\prod k_i!(n_i)^{k_i}$ sets of products in the integrand of (44) leading to one such integral, so that $H_{\nu m}\{\mathfrak{v}\}v^{-m}$ may be written as

$$H_{\nu m}\{\mathfrak{v}\}v^{-m} = \sum \{\{\mathfrak{y}_i\}_\nu\}_c \sum_{\substack{n_i \\ \sum_i n_i k_i = m}} \sum_{k_i} \prod_i \frac{[h_{\mu_i n_i}\{\mathfrak{y}_i\}_\nu v^{-n_i}]^{k_i}}{k_i!} \tag{54}$$

and

$$\begin{aligned} G_\nu\{\mathfrak{v}\} &= \sum_{m \geq 1} v^{-m} H_{\nu m}\{\mathfrak{v}\} \\ &= \sum \{\{\mathfrak{y}_i\}_\nu\}_c \prod_i \left[\sum_{k \geq 1} \frac{1}{k!}\left(\sum_{n \geq 1} h_{\mu_i n}\{\mathfrak{y}_i\}_\nu v^{-n}\right)^k\right] \\ &= \sum \{\{\mathfrak{y}_i\}_\nu\}_c \prod_i [(\exp \sum_{n \geq 1} h_{\mu_i n}\{\mathfrak{y}_i\}_\nu v^{-n}) - 1]. \end{aligned} \tag{55}$$

It is seen, then, from (40) and (55) that

$$\ln F_\lambda\{\mathfrak{\lambda}\} = \sum_{\lambda \geq i > j \geq 1} -\frac{u(\mathbf{i}, \mathbf{j})}{kT} + \sum \{\mathfrak{y}_i\}_\lambda [\sum_{n \geq 1} h_{\mu_i n}\{\mathfrak{y}_i\}_\lambda v^{-n}]. \tag{56}$$

By comparison with Eq. (11) one finds that

$$h_2(1, 2) = -\frac{u(1, 2)}{kT} + \sum_{n \geq 1} v^{-n} h_{2,n}(1, 2), \tag{57}$$

$$h_\mu\{\mathfrak{y}\} = \sum_{n \geq 1} v^{-n} h_{\mu n}\{\mathfrak{y}\}, \tag{57'}$$

as equations for the auxiliary functions defined previously.

It may be well to mention at this place that if the functions $g_\nu\{\mathfrak{v}\}$, defined by Eqs. (7) or (8), are formed, the integral $\int g_\nu\{\mathfrak{v}\}d\tau_i$ of Eq. (10) is *not* zero. This is due to the fact that Eq. (38) is not correct for terms as small as $1/N$. For instance:

$$\begin{aligned} g_2(1, 2) &= F_2(1, 2) - 1 \\ &= [1 + f(1, 2)] \sum_{l \geq 1} B_{2l}(1, 2) v^2 Z^{l+2} + f(1, 2), \end{aligned} \tag{58}$$

$$\int g_2(1, 2) d\tau_2 = v[\sum_{l \geq 1} l v b_l Z^l - 1] \neq 0. \tag{58'}$$

IV. Functions in Condensed Systems

Equation for the distribution functions, $F_\lambda\{\mathfrak{\lambda}\}$

In condensed systems, for which the developments of the last section diverge, Eq. (36′) may be used as a starting point instead of the completely expanded form (36).

Consider a single product, $\prod f_{\kappa i}^{\nu_{\kappa i}}$, of the sum of products occurring in the integrand of (36′), in which f functions to only m molecules of some subset $\{\mathbf{m}\}_{N-m}$ of the set $\{\mathbf{N}-\boldsymbol{\lambda}\}$ are present with nonzero coefficients $\nu_{\kappa i}$. The integral over the remaining $N-m-\lambda$ molecules of the set $\{\mathbf{N}-\boldsymbol{\lambda}\}$ leads to

$$\frac{V^{\lambda}}{\mathbf{Q}_{\tau}(N, V, T)} \int\int \cdots \int [\exp(-U\{\mathbf{N}-\boldsymbol{\lambda}\}/kT)] d\tau\{\mathbf{N}-\mathbf{m}-\boldsymbol{\lambda}\} = V^{\lambda-m} \frac{\mathbf{Q}_{\tau}(N-\lambda, V, T)}{\mathbf{Q}_{\tau}(N, V, T)} F_m\{\mathbf{m}\} = v^{\lambda-m} \frac{1}{N^m} Z^{\lambda} F_m\{\mathbf{m}\},$$

from Eqs. (32) and (34).

Since there are $(N-\lambda)!/(N-\lambda-m)!m! \cong N^m/m!$, (for $\lambda, m \ll N$), different sets of m molecules each, leading to the same integrals, one may write

$$F_{\lambda}\{\boldsymbol{\lambda}\} = (vZ)^{\lambda}[\exp(-U\{\boldsymbol{\lambda}\}/kT)] \sum_{m\geq 0} \frac{1}{m!} \frac{1}{v^m} \int\int \cdots \int \sum_{\substack{\nu_{\kappa i}=0,1 \\ \sum_i \nu_{\kappa i} \geq 1}} \prod_{\substack{\lambda \geq \kappa \geq 1 \\ m \geq l \geq 1}} f_{\kappa i}^{\nu_{\kappa i}} F_m\{\mathbf{m}\} d\tau\{\mathbf{m}\}, \quad (59)$$

in which the first term of the sum, for which m is zero, is to be taken equal to unity.

The rather complicated conditions imposed upon the product of $f_{\kappa i}$'s for given values of m are such that the sum of products is just

$$\prod_i [(\exp \sum_{\kappa=1}^{\kappa=\lambda} -u(\kappa, \mathbf{i})/kT) - 1].$$

The alternative form for the expression (59),

$$F_{\lambda}\{\boldsymbol{\lambda}\} = (vZ)^{\lambda}[\exp(-U\{\boldsymbol{\lambda}\}/kT)] \sum_{m\geq 0} \frac{1}{m!} \frac{1}{v^m} \int\int \cdots \int \prod_{i=1}^{i=m} [(\exp \sum_{\kappa=1}^{\kappa=\lambda} -u(\kappa, \mathbf{i})/kT) - 1] F_m\{\mathbf{m}\} d\tau\{\mathbf{m}\} \quad (59')$$

is, for some purposes, more convenient.

It is to be noted that for a system composed of molecules which repel at short distances the functions $F_m\{\mathbf{m}\}$ become zero for large m values if the coordinates of all the m molecules are limited to lie in one or several comparatively small regions of space. Since the product of $f_{\kappa i}$ functions in the integrand becomes zero if all the molecules of the set $\{\mathbf{m}\}$ do not lie in one or other of the small regions close to the positions of the molecules of the set $\{\boldsymbol{\lambda}\}$, it follows that the terms of sufficiently large m in the sum are zero. The series in the Eqs. (59) or (59′) necessarily converge for finite values of λ, for all types of physical systems.

Equation for the fugacity, Z

The special case of Eq. (59) that $\lambda=1$, $F_1(1)=1$, $U(1)=0$, leads to an expression for the fugacity, Z, of the system:

$$\frac{1}{(vZ)} = \sum_{m\geq 0} \frac{1}{m!} \frac{1}{v^m} \times \int\int \cdots \int \prod_{m\geq i\geq 1} f(\mathbf{0}, \mathbf{i}) F_m\{\mathbf{m}\} d\tau\{\mathbf{m}\}. \quad (60)$$

With Eq. (31) this leads to an expression for the chemical potential, μ, since

$$\frac{1}{(vZ)} = \left(\frac{h^2}{2\pi mkT}\right)^{\frac{3}{2}} \frac{1}{v} e^{-\mu/kT} \quad (61)$$

It is tempting, and the authors had thought it possible, to use Eq. (60) alone, with certain geometrical conditions on the $F_m\{\mathbf{m}\}$, to determine both the fugacity and the distribution functions by means of a minimum condition. If Eq. (60) were interpreted as giving the free energy of a system restricted to that part of the configuration space defined by any arbitrary set of geometrically possible distribution functions, then the set of $F_m\{\mathbf{m}\}$ functions of the equilibrium configuration would be those which led to

a minimum of the free energy, or a minimum of the fugacity Z.

Actually Eq. (60) gives the chemical potential, or differential coefficient of the free energy, upon the addition of one molecule to a system which has *initially* the distribution of molecules given by the $F_m\{\mathfrak{m}\}$, but does not necessitate that the molecular distribution functions remain unchanged upon the addition of that and other molecules to the system. As a result the calculated μ is *not* a true differential coefficient for the free energy, except for the equilibrium distribution. For instance the assumption that all $F_m\{\mathfrak{m}\}$'s are unity, which is clearly a geometrically possible distribution, leads to the simple equation

$$\frac{1}{(vZ)} = \sum_{m\geq 0} \frac{1}{m!}\frac{\beta_1{}^m}{v^m} = e^{\beta_1/v}, \tag{62}$$

in which

$$\beta_1 = 4\pi\int_0^\infty r^2 f(r)\,dr. \tag{63}$$

This equation predicts a lower free energy than that of the equilibrium distribution at volumes and temperatures for which the liquid or solid phase is stable. It is, however, the correct first approximation to the free energy of a dilute gas, and is the same as the first approximation derived out of Eq. (50).

Equation for the auxiliary functions, $h_\mu\{\mathfrak{y}\}$

The sum of products occurring under the integral sign for each value of m in Eq. (59),

$$\sum_{\nu_{\kappa i}=0,\,1}{}^{(\Sigma_i \nu_{\kappa i}\geq 1)} \prod_{\substack{\lambda\geq k\geq 1\\ m\geq 1\geq 1}} f_{\kappa i}{}^{\nu_{\kappa i}},$$

may be separated as a sum of terms, each involving only coordinates of some subset $\{\mathfrak{y}\}_\lambda$ of the set $\{\mathfrak{z}\}$,

$$\sum\{\mathfrak{y}\}_\lambda \sum_{\nu_{\kappa i}=0,\,1}{}^{(\Sigma_\kappa \nu_{\kappa i}\geq 1,\ \Sigma_i \nu_{\kappa i}\geq 1)} \prod_{\substack{\mu\geq\kappa\geq 1\\ m\geq i\geq 1}} f_{\kappa i}{}^{\nu_{\kappa i}}.$$

Using also Eq. (7) for the development of the $F_m\{\mathfrak{m}\}$ in sums of products of the functions $g_\nu\{\mathfrak{v}\}$, and by the use of the integrals

$$h_{\mu\nu}{}^*\{\mathfrak{y}\} = \frac{1}{\nu!}\int\int\cdots\int \sum_{\nu_{\kappa i}=0,\,1}{}^{(\Sigma_\kappa \nu_{\kappa i}\geq 1,\ \Sigma_i \nu_{\kappa i}\geq 1)} \prod_{\substack{\mu\geq\kappa\geq 1\\ \nu\geq i\geq 1}} f_{\kappa i}{}^{\nu_{\kappa i}} g_\nu\{\mathfrak{v}\}\,d\tau\{\mathfrak{v}\}, \tag{64}$$

one may express (59) as

$$F_\lambda\{\mathfrak{z}\} = (vZ)^\lambda[\exp(-U\{\mathfrak{z}\}/kT)]\sum_{m\geq 0}\frac{1}{v^m}\sum_{n_\nu\geq 0}\prod_\nu \frac{[\sum\{\mathfrak{y}\}_\lambda\, h_{\mu\nu}{}^*\{\mathfrak{y}\}_\lambda]^{n_\nu}}{n_\nu!}$$

$$= (vZ)^\lambda \exp\left[-U\{\mathfrak{z}\}/kT + \sum\{\mathfrak{y}\}_\lambda \sum_{\nu\geq 1} v^{-\nu} h_{\mu\nu}{}^*\{\mathfrak{y}\}_\lambda\right]. \tag{65}$$

Since

$$F_1(1) = 1 = vZ\exp\left[\sum_{\nu\geq 1} v^{-\nu} h_{1\nu}{}^*(1)\right], \tag{65'}$$

and λ different sets $\sum_{\nu\geq 1} v^{-\nu} h_{1\nu}{}^*\{\kappa\}$ occur in the exponent for $F_\lambda\{\mathfrak{z}\}$, these terms exactly cancel $(vZ)^\lambda$. The Eq. (65) then leads to

$$h_2(1,2) = -u(1,2)/kT + \sum_{\nu\geq 1} v^{-\nu} h_{2\nu}{}^*(1,2), \tag{66}$$

$$h_\mu\{\mathfrak{y}\} = \sum_{\nu\geq 1} v^{-\nu} h_{\mu\nu}{}^*\{\mathfrak{y}\}, \tag{66'}$$

where the functions $h_\mu\{\mathfrak{y}\}$ are those defined by Eq. (11), and calculated for the gas in Eq. (57) and (57').

It is to be noted that whereas the functions $h_{\nu n}\{\mathfrak{v}\}$ defined by Eq. (53), and occurring in (57) and (57') are (for reasonable values of n at least) independent of the volume, the functions $h_{\mu\nu}{}^*\{\mathfrak{y}\}$ are integrals over the volume dependent functions $g_\nu\{\mathfrak{v}\}$, and are consequently themselves functions of the volume v. Equations (57) and (57') then give a true development of the auxiliary functions

$h_\mu\{\mathfrak{y}\}$ as a power series in the inverse volume per molecule, but (66) and (66′) do not. It is consequently impossible to identify the coefficients in the two sets of equations, and $h_{\mu n}\{\mathfrak{y}\} \neq h_{\mu n}{}^*\{\mathfrak{y}\}$.

Equation (65) which may be written as

$$\ln Z = -\ln v - \sum_{\nu \geq 1} h_{1\nu}{}^* v^{-\nu} \tag{67}$$

resembles Eq. (50), which formed the expression for calculation of the fugacity Z as a power series of the inverse volume, in the theory of the imperfect gas. However, the quantities $h_{1\nu}{}^*$, which are not volume independent, are not to be identified with the coefficients, β_k, of the power series in v^{-k}.

Equations using the population probability functions, $P(\{\mathbf{R}\}\{\nu\})$

Suppose that the three-dimensional space within a distance d or less of any of a set $\{\lambda\}$ of λ points is divided into R small non-overlapping contiguous regions that completely fill this part of space. The running index r, $R \geq r \geq 1$, will be used to indicate the region and $\mathbf{r}$ to indicate the position of the region. The symbol $\{\mathbf{R}\}$ will be used for the set of regions. It will be assumed that the regions are chosen sufficiently small so that the function $f(\boldsymbol{\kappa}, \mathbf{r})$ of the distance between the point $\boldsymbol{\kappa}$ of the set $\{\lambda\}$ and the position $\mathbf{r}$ of the rth region is essentially constant for all parts of the region. It will also be assumed that the distance d has been chosen sufficiently great so that $f(\boldsymbol{\kappa}, \mathbf{i})$ to any point $\mathbf{i}$ not in one of the R regions is essentially zero.

Then any integral of the type occurring in (59′) may be written as

$$\int\int\cdots\int \frac{1}{m!}\frac{1}{v^m} \prod_{m \geq i \geq 1} [(\exp \sum_{\kappa=1}^{\kappa=\lambda} -u(\boldsymbol{\kappa},\mathbf{i})/kT) - 1] F_m\{\mathbf{m}\} d\tau\{\mathbf{m}\}$$

$$= \sum_{\nu_r} {}^{(\sum_r \nu_r = m)} \prod_{R \geq r \geq 1} [(\exp \sum_{\kappa=1}^{\kappa=\lambda} -u(\boldsymbol{\kappa}, \mathbf{r})/kT) - 1]^{\nu_r} \phi(\{\mathbf{R}\}\{\nu\}),$$

the ϕ functions of which are those defined by Eq. (21), (23) and (24).

The summation over m lifts the restriction on the values of the ν_r's, so that

$$F_\lambda\{\lambda\} = (vZ)^\lambda [\exp(-U\{\lambda\}/kT)] \sum_{\nu_r \geq 0} \prod_{R \geq r \geq 1} [(\exp \sum_{\kappa=1}^{\kappa=\lambda} -u(\boldsymbol{\kappa}, \mathbf{r})/kT) - 1]^{\nu_r} \phi(\{\mathbf{R}\}\{\nu\}), \tag{68}$$

in which the first term of the sum, with all ν_r's equal to zero, is unity.

If one now forms the sum

$$\sum_{\mu_r \geq 0} [\exp (\sum_{\kappa=1}^{\kappa=\lambda} \sum_{r=1}^{r=R} -\mu_r u(\boldsymbol{\kappa}, \mathbf{r})/kT)] P(\{\mathbf{R}\}\{\mathfrak{y}\})$$

and inserts Eq. (25) for $P(\{\mathbf{R}\}\{\mathfrak{y}\})$, one finds

$$\sum_{\mu_r \geq 0} \{ \prod_{R \geq r \geq 1} \exp (\sum_{\kappa=1}^{\kappa=\lambda} -\mu_r u(\boldsymbol{\kappa}, \mathbf{r})/kT) \} \sum_{\nu_r \geq \mu_r} \prod_{R=r=1} \frac{\nu_r!(-1)^{\nu_r - \mu_r}}{(\nu_r - \mu_r)!\nu_r!} \phi(\{\mathbf{R}\}\{\nu\})$$

$$= \sum_{\mu_r \geq 0} \prod_{R \geq r \geq 1} \left\{ \sum_{\mu_r=1}^{\mu_r=\nu_r} \frac{\nu_r!(-)^{\nu_r-\mu_r}}{(\nu_r-\mu_r)!\mu_r!} [\exp \sum_{\kappa=1}^{\kappa=\lambda} -u(\boldsymbol{\kappa}, \mathbf{r})/kT]^{\mu_r} \right\} \phi(\{\mathbf{R}\}\{\nu\})$$

$$= \sum_{\nu_r \geq 0} \prod_{R \geq r \geq 1} [(\exp \sum_{\kappa=1}^{\kappa=\lambda} -u(\boldsymbol{\kappa}, \mathbf{r})/kT) - 1]^{\nu_r} \phi(\{\mathbf{R}\}\{\nu\}),$$

so that

$$F_\lambda\{\lambda\} = (vZ)^\lambda [\exp(-U\{\lambda\}/kT)] \sum_{\mu_r \geq 0} \{\exp \sum_{\kappa=1}^{\kappa=\lambda} \sum_{r=1}^{r=R} -\mu_r u(\boldsymbol{\kappa}, \mathbf{r})/kT\} P\{\mathbf{R}\}\{\mathfrak{y}\}). \tag{69}$$

The special case $\lambda=1$ yields an equation for the fugacity:

$$\frac{1}{(vZ)}=\left(\frac{h^2}{2\pi mkT}\right)^{\frac{3}{2}}\frac{1}{v}e^{-\mu/kT}=\sum_{\mu_r\geq 0}\{\exp\sum_{r=1}^{r=R}-\mu_r u(\mathbf{0},\mathbf{r})/kT\}P(\{\mathbf{R}\}\{\boldsymbol{\mu}\}). \tag{70}$$

Equation (70) is essentially that which forms the basis of the cell method of computing liquid properties[4] and which is so frequently used. Thus the direct connection between the methods used for the imperfect gas, and the cell method, has been demonstrated.

It is worthy of remark that the population probability function $P(\{\mathbf{R}\}\{\boldsymbol{\mu}\})$ appearing in (70) gives the probability that the set of cells $\{\mathbf{R}\}$ shall be occupied by the numbers of molecules represented by the set $\{\boldsymbol{\mu}\}$, if the origin is *not* occupied by a molecule. This fact has not always been adequately recognized, although explicitly stated and handled by Kirkwood.

V. Numerical Evaluations

It has been the purpose of this paper to derive formal equations, independently of special numerical assumptions, for the various functions which may be used to describe the distribution of molecules in a system. Nevertheless a few remarks about the numerical possibilities of the equations derived here may somewhat relieve the apparent sterility of the formal mathematics.

Equations (57) and (57′), for the gas, may be used by direct numerical integration. The first few terms in the power series offer no exceptional computational difficulties. The evaluation is greatly facilitated by the use of the Fourier transform, in three-dimensional space, of the function $f(r)$, and with the use of this transformation part of the summation can be performed before integration.

The set of Eqs. (59) or (59′) must be satisfied by the equilibrium distribution functions of any system. Although for given finite λ the summation on the right always converges, the set, in a sense, diverges, since for a given value of λ on the left the summation on the right always includes values of m greater than λ. This alone does not preclude the use of the set for approximate numerical evaluation, since presumably a knowledge of $F_\lambda\{\boldsymbol{\lambda}\}$ for a given moderately large value of λ permits estimation of the function for larger values of λ. It is for this reason that the auxiliary functions $h_\mu\{\boldsymbol{\mu}\}$ have been introduced, since presumably these decrease in numerical importance as μ increases, thus giving a method of estimation of distribution functions for large values of λ from a knowledge of those with small λ.

For instance a zeroth order approximation, $h_2(1,2)=-u(1,2)/kT$, $h_\mu\{\boldsymbol{\mu}\}=0$, $\mu\geq 3$, may be used with (13) to calculate zeroth order approximations of the $g_\nu\{\boldsymbol{\nu}\}$ functions. The use of these in (64), (66) and (66′) lead to first-order approximations for the $h_\mu\{\boldsymbol{\mu}\}$'s, which may again be fed into (13) to obtain better $g_\nu\{\boldsymbol{\nu}\}$'s. The repetition of this process is limited only by patience, which would, however, be a severe limitation, since the integrations involved in (64) are extremely tedious.

The repetition of this process, and subsequent development of the $h_\mu\{\boldsymbol{\mu}\}$'s in a power series of v^{-1} will lead to (57) and (57′) which diverge for condensed systems, which does not, however, preclude that the method without development in a power series would lead to a convergent result.

Equations (69) and (70), however, appear to lead most easily to useful numerical evaluations. The cells or regions may be chosen of uniform volume v_0, arranged in, say, the cubic close packed lattice, and the volume v_0 so chosen that the "diameter" d, of the cells is just the largest distance for which the exponential of $-u(d)/kT$ may be regarded as zero. The equations predict in this case that the probability of having two or more molecules in a cell is zero. The summations are then restricted to $\mu_r=0$ and 1.

As a zeroth-order approximation one may use the assumption that the probabilities of population of the cells are independent of each other,

[4] See for instance E. A. Guggenheim, Proc. Roy. Soc., London, **A135**, 181 (1932).

in which case the probability that a cell be occupied is v_0/v, and that it be empty is $(v-v_0)/v$. The use of this zeroth-order assumption in (70) leads to a simple expression for the fugacity Z, which is of the correct qualitative form, and is essentially that usually derived by the use of the cell method.

This same zeroth-order approximation may, however, be used in (69) to calculate any of the distribution functions $F_\lambda\{\lambda\}$, and these values, with the use of (21) and (25) may be used to arrive at closer approximations to the correct population probability functions, $P(\{\mathbf{R}\}\{\boldsymbol{\mu}\})$.

Although this method appears to the authors to be definitely less elegant and less satisfactory than operation with the continuous distribution functions, and the method of integration, it appears to offer certain advantages of numerical simplicity. At least the equations offer a possibility of consistently working the usual cell method of approach to higher approximations of accuracy.

JANUARY, 1941 JOURNAL OF CHEMICAL PHYSICS VOLUME 9

Theory of the Transition in KH_2PO_4*

J. C. SLATER
Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received October 10, 1940)

Potassium dihydrogen phosphate contains phosphate groups connected by hydrogen bonds. Different possible arrangements of the hydrogens result effectively in different orientations of the $(H_2PO_4)^-$ dipoles. Since these have the lowest energy when pointing along the axis of the crystal, there is a tendency toward spontaneous polarization along this axis, resulting in the well-known transition, similar to Rochelle salt, with polarization below the Curie point. The theory of this transition is worked out, using statistical methods to count the number of arrangements of hydrogens consistent with each total polarization of the crystal, and deriving the free energy. It is found that the theory predicts a phase change of the first order, with sudden transition from the polarized state at low temperature to the unpolarized state at high temperature, rather than the lambda-point transition or phase change of the second order which is observed. However, the observed transition is confined to a very narrow temperature range compared to that predicted, for instance, by the Weiss theory, so that it seems as if it might be merely a broadened transition of the first order. It is suggested that the broadening may result from the irregular shifts of transition temperatures of individual domains in the crystal on account of stresses resulting from the large piezoelectric effect and the resulting deformation of the crystal below the transition point. The susceptibility above the Curie point comes out by the theory to be 4.33 times as great as it should according to the Weiss theory, a result which seems to be in general agreement with experiment. The entropy change in the transition is given by the theory as 0.69 unit, somewhat smaller than the observed value of about 0.8 unit. No explanation is suggested for this discrepancy.

IT has been known for several years[1] that potassium dihydrogen phosphate, KH_2PO_4, has a phase change of the second order at about 122°K, showing a dielectric analog of ferromagnetism, as does Rochelle salt, below this temperature. The crystal structure of Rochelle salt has, unfortunately, not been worked out until very recently, so that theories of its dielectric behavior have been handicapped, but potassium phosphate has a simple crystal structure,[2] and one can proceed with a good deal of assurance in explaining its properties. It is the purpose of this paper to work out the basic statistical mechanical theory of the phase change, and to show how far the theory agrees with experiment. A number of points in the calculation prove to be of general interest, not merely

* Presented at the Symposium on the Structure of Molecules and Aggregates of Molecules at the Fifth Annual Symposium of the Division of Physical and Inorganic Chemistry of the American Chemical Society, Columbia University, New York, December 30, 1940 to January 1, 1941.

[1] G. Busch and P. Scherrer, Naturwiss. **23**, 737 (1935); G. Busch, Helv. Phys. Acta **11**, 269 (1938); C. C. Stephenson and J. G. Hooley, Phys. Rev. **56**, 121 (1939); W. Bantle and P. Scherrer, Nature **143**, 980 (1939); Mendelssohn, Nature **144**, 595 (1939).

[2] J. West, Zeits. f. Krist. **74**, 306 (1930).

applying to this particular problem. For one thing, the statistical methods used are somewhat unusual, and might be applicable to other problems. And in addition, the results of the statistics are unusual. At first sight, it might be thought that the problem was closely analogous to the Weiss theory of ferromagnetism. It proves not to be, however. The dielectric susceptibility above the Curie point is given by a different formula from the Weiss theory. And below the Curie point, instead of predicting the rather gradual increase of saturation polarization with decreasing temperature which the Weiss theory gives, the present calculation indicates an instantaneous change, or a phase change of the first order, with complete polarization at all temperatures below the Curie point, and with a latent heat instead of a specific heat anomaly. This of course is not in agreement with experiment, and we are forced to suggest tentative explanations for the discrepancy. Nevertheless, the observed specific heat anomaly is much sharper than that given by the Weiss theory, and is in fact not far from a phase change of the first order. And it is very interesting that a theory, based as this one is on perfectly straightforward statistical methods, can give such a sharp transition, since most observed lambda-points are actually much sharper than the Weiss theory suggests.

1. The Structure of the Crystal

The elements of the crystal structure which interest us are the phosphate groups, $(PO_4)^{-3}$, and the hydrogen ions; the potassiums merely have fixed positions in the crystal, and serve to hold it together. A phosphate group consists of a phosphorus tetrahedrally surrounded by four oxygens. The structure of the crystal is shown in Fig. 1. It is seen that the tetrahedra are arranged in layers, with a square arrangement in each layer. Each phosphate group is surrounded tetrahedrally by four other phosphate groups; this is seen particularly easily from the group in the center of the unit cell shown in Fig. 1. The position of the hydrogens is not shown by x-ray analysis; but symmetry demands that one be located somewhere on the connecting line between each pair of phosphate groups, as shown in the figure. It is most likely, however, that instead of being midway between the two phosphates, as shown in the figure, each hydrogen is near one or the other of its neighboring phosphates. On the average, each phosphate is likely to have two of its neighboring hydrogens close to it, forming a group $(H_2PO_4)^-$, the other two neighbors going with other phosphate groups. In fact, we may imagine that it is only very rarely that a phosphate group has any number of neighboring hydrogens other than two; in other words, we may suppose that the pair of ions $(HPO_4)^{--}$, (H_3PO_4), which would be formed by transfer of an extra hydrogen from one $(H_2PO_4)^-$ group to its neighbor, would have a sufficiently high energy so that the Boltzmann factor would make them very unlikely to occur, at ordinary temperatures.

It will be seen that this situation is very much like what occurs in the case of ice,[3] or, presumably, in liquid water. There each oxygen is tetrahedrally surrounded by four others, with

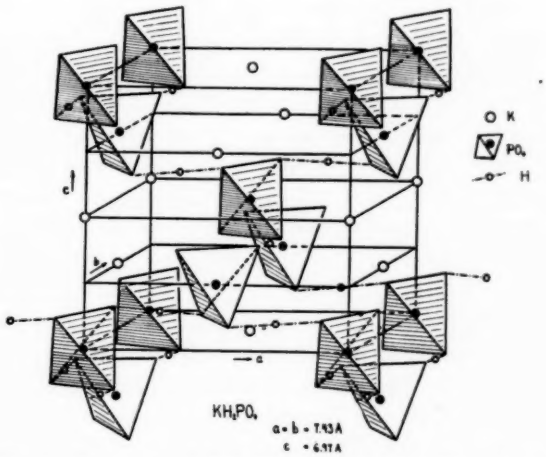


Fig. 1. Structure of KH_2PO_4. Figure from Dr. C. C. Stephenson, by permission.

a hydrogen along the line joining each oxygen to its neighbors, but practically always two of the four hydrogens are near a given oxygen, and the other two near neighboring oxygens. In other words, practically always we have neutral H_2O molecules, the ions $(HO)^-$ and $(H_3O)^+$ being very rare, as is well known from electrochemical data. And as in ice, there are many different arrangements of the hydrogens consistent with

[3] L. Pauling, J. Am. Chem. Soc. **57**, 2680 (1935).

the structure. We remind ourselves of Pauling's[3] argument giving the number of arrangements. If there are N oxygens (in ice) or N phosphate groups (in KH_2PO_4), and $2N$ hydrogens, then if the hydrogens were independent of each other, each would have two possible positions, so that there would be 2^{2N} different configurations. But now consider one water molecule, or $(H_2PO_4)^-$ group. It has four bonds to neighbors, so that the hydrogens surrounding it can be arranged in $2^4=16$ ways. Of these 16, however, only 6 arrangements ($=(4\times3)/2$) correspond to having two of the hydrogens near the group, the others away from it. Thus only $(6/16)2^{2N}$ of the arrangements are allowed, as far as this particular group is concerned. If we can treat all the groups as independent, each of them will allow a fraction (6/16) of all arrangements. Thus the total number of allowed arrangements will be $(6/16)^N 2^{2N}=(3/2)^N$. The assumption that the groups can be treated independently, and their factors (6/16) simply multiplied, does not seem very sound at first sight; but more careful examination verifies the correctness of Pauling's calculation. We remember how Pauling used this calculation: He noted that all of the $(3/2)^N$ states had the same energy, so that they all persisted to the absolute zero, and as a result it is to be expected, as is actually found experimentally, that ice at the absolute zero has an entropy $k \ln (3/2)^N = Nk \ln (3/2)$, rather than zero.

While there is a great similarity between our problem and that of ice, there is also a great difference. In ice, all six allowed arrangements of the hydrogens about an oxygen are equivalent crystallographically, and all correspond to the same energy. In KH_2PO_4, however, two of the arrangements are different from the other four, and can be expected to have a different energy. The two different arrangements are those in which the two hydrogens near a PO_4 group are either both on the upper side of the group (that is, the side toward the positive c direction, as in Fig. 1) or both on the lower side; the four other arrangements are those in which one hydrogen is up, one down. Each arrangement, we note, makes the $(H_2PO_4)^-$ group an electrical dipole. In the arrangement with the hydrogens up, this dipole points along the positive c direction; in that with the hydrogens down, it points along the negative c; while in the other four arrangements, with one hydrogen up and one down, it points in a direction in the plane perpendicular to c. Since the c axis is a preferred axis in the crystal, the two orientations of the dipole, along $\pm c$, can have different energy from those perpendicular to c. The dielectric behavior of the crystal shows that the orientations along c or $-c$ must have a lower energy than those at right angles. Thus there is a tendency for a spontaneous polarization, at low temperatures, along these directions. The state of lowest energy will come when each dipole is oriented either along c or $-c$. A little reflection will show that in this case either all dipoles must be along $+c$, or all along $-c$; for any arrangement in which some were along $+c$, some along $-c$, would demand having some ions with the wrong number of hydrogens. Thus at the absolute zero, the crystal will be completely polarized, on account of the low energy of this state; while at high temperatures, the dipoles will be arranged at random, on account of the resulting high entropy. It is the purpose of this paper to investigate the statistical mechanics of this transition. We note the contrast to the case of ice, where the disordered state persists to the absolute zero, simply because a state with an outstanding dipole moment does not have a lower energy, so that ice has no Curie point.

It should be clearly understood that there is a well-known point of view regarding ice, and other materials containing dipoles, which is in contradiction to this theory.[4] This point of view is based on the Lorentz factor $4\pi/3$ occurring in the theory of dielectrics, and its relation to the Weiss theory of ferromagnetism. According to that theory, in a crystal with a simple cubic arrangement of parallel dipoles, there is an electric field at any dipole which is greater by $(4\pi/3)P$, where P is the polarization vector, than the average field throughout the crystal. This added field results in a lower energy if the dipole in question points parallel to the polarization vector, than we should have if it were differently oriented. The result is a term in the internal energy which is lower for the polarized crystal

[4] For a general discussion of this point, see J. H. Van Vleck, J. Chem. Phys. 5, 556 (1937).

than for a crystal in which the dipoles of all molecules are oriented at random, so that there is no polarization. We expect a phase change at low temperatures, then, from the high temperature, nonpolarized state to a low temperature, polarized state. When investigated, this proves to be a phase change of the second order, occurring gradually through a long range of temperature, and it is analogous to the Weiss theory of ferromagnetism, except that in the latter theory another factor must be used in place of $4\pi/3$. The magnitudes are such that if the theory were correct, ordinary materials having dipole molecules, such as ice, would have Curie points, and spontaneous electrification, and the Curie points in ordinary cases would come at high temperatures, of the order of several thousand degrees. This situation has been called "the $4\pi/3$ catastrophe." It should be understood clearly how we avoid this catastrophe, and how our theory of the spontaneous polarization of KH_2PO_4 is completely different from this theory. According to the $4\pi/3$ theory, the field at a dipole is $(4\pi/3)P$, so that the energy of this dipole is lowest when it points along the direction of P. In our theory, however, the energy of a given dipole (H_2O, or $H_2PO_4^-$, in the two cases), depends not on the external polarization, but on its neighbors. We can consider each dipole to have six possible orientations; but if the hydrogen ions of the neighboring groups are fixed in position, only one of these orientations has a low energy, the one in which the two hydrogens of the dipole in question point to neighbors which do not have other hydrogens in the way. In other words, for low energy, there can be but one hydrogen ion on each bond joining neighboring groups. This means that the lowest energy of a dipole comes in different directions for different dipoles, but this energy is so much lower than for any other orientation that we specifically neglect the possibility of any other orientation, whether we are above or below the Curie point. At all temperatures which we consider, each dipole is oriented in the most favorable way possible, with respect to its neighbors. But on account of the form of the dipoles (a central group with two hydrogens attached) this does not by any means imply a net polarization of the whole crystal. In a sense, all the states we consider are below the "Curie point" of the $4\pi/3$ catastrophe; to get above that point, we should have to have each dipole oriented independently of its immediate neighbors, which would mean the frequent existence of two, or no, hydrogens on the bond between two groups. The energy which would come from the $4\pi/3$ term, the interaction energy of the dipoles, is present in a latent way in our theory, as a constant, independent of the particular arrangement of dipoles which we consider, since all arrangements are equivalent as far as that energy is concerned.

On account of the geometry of the crystal, however, we have our $(3/2)^N$ possible states, all of the same energy as far as this dipole-dipole interaction is concerned, and yet having different outstanding dipole moments, from zero up to a maximum. It will be our task to find how many of these states correspond to each particular dipole moment. There is no tendency to prefer states of large dipole moment, on account of dipole-dipole interaction, so that in the case of ice there is no tendency to a Curie point. And in KH_2PO_4 also, the dipole-dipole interaction does not favor states of large dipole moment. In that case, however, as we have seen, each individual dipole has a lower energy if it is oriented parallel or opposite to the c direction, and this brings about a tendency to spontaneous polarization, not because the dipoles want to be parallel to each other, but because they prefer this definite direction with respect to the crystal. It is for this reason that the spontaneous polarization is always found along the c axis.

The explanation we have given of the lack of the $4\pi/3$ catastrophe, and the lack of a Curie point for ice, is not original with us. It is essentially similar to the argument given by Onsager,[5] who considers that in dipole liquids each dipole is located in a field whose orientation depends largely on the orientation of that dipole itself, so that although the energy term connected with the dipole-dipole interaction is large, it does not depend on the dipole moment as a whole. The method used by Onsager is rather different from the present one, however, so that it has seemed worth while to state the argument at this point. The specific application to KH_2PO_4, by a

[5] L. Onsager, J. Am. Chem. Soc. **58**, 1486 (1936).

method somewhat similar to that used in this paper, has also been given by Onsager.[6] Nevertheless, the present treatment is different from Onsager's, and more complete in some respects, and for that reason it seems worth presenting, though the results agree with Onsager's as far as his go.

2. Elementary Theory of the Transformation

First we shall present an elementary theory of the potassium phosphate transformation, leading to essentially the same results as Onsager's[6] theory. We shall find a number of points about this theory that are not entirely satisfactory; but it is simple and easy to understand and gives results which agree with the more complicated theory which we present later, so that it seems worth presenting. The limitations of this simple theory are two. In the first place, we consider only two states of the crystal: the perfectly ordered state, in which all dipoles point in the same direction, and the completely disordered state. We do not consider intermediate states. The ordered state has the maximum possible dipole moment, and polarization, while the disordered state has no net polarization. Thus we are bound to get a definite equilibrium temperature between the states, with polarization below this temperature and not above, and a phase change of the first order, with latent heat. The second limitation is that we make the enumeration of configurations, which we need in computing the partition function of statistical mechanics, in a somewhat unsatisfactory way. Both these shortcomings are overcome in the later treatment.

Let us assume that there are N phosphate groups, and $2N$ hydrogens, in the crystal. We shall assume that when the hydrogens about a phosphate group are arranged so that the dipole points along the $\pm c$ direction, its energy is zero (choosing the arbitrary additive constant to secure this), and when it is at right angles to this direction, the energy is ϵ. We shall now compute the partition function Z, the sum over all states of the quantity $\exp(-\text{energy}/kT)$. To do this, we must find how many states there are with each possible value of the energy. Plainly the energy can take on values zero (when all dipoles are along $\pm c$), ϵ (when one dipole is perpendicular), 2ϵ (when two dipoles are perpendicular), etc., up to $N\epsilon$ (when all dipoles are perpendicular to c). We use a method of enumeration similar to Pauling's to find the number of arrangements with each such energy. We shall point out shortcomings of this method later.

Let us first find the number of arrangements with energy of zero. As in Pauling's argument, outlined in Section 1, we assume that there are 2^{2N} arrangements of the hydrogens, of which a great many are excluded because they correspond to an incorrect number of hydrogens on some phosphate groups, and many more are now excluded because they correspond to some dipoles pointing perpendicular to the c axis. In fact, the chance that a given dipole point along the $+c$ direction is 1/16 (that is, of the 16 ways of arranging the four hydrogens surrounding a phosphate group, one corresponds to orientation along $+c$), and the chance that it point along $-c$ is likewise 1/16, so that the chance that it point along one or the other of these directions is the sum of the probabilities of the two mutually exclusive events, or 2/16. Similarly the chance that each other dipole point along the $\pm c$ is 2/16. If then we follow Pauling, we conclude that the chance that each dipole point along either $+c$ or $-c$ is $(2/16)^N$. Hence the number of arrangements of hydrogens consistent with this energy should be $2^{2N}(2/16)^N$.

(We note that this result is certainly wrong, warning us at the beginning to be on our guard against this method of computing probabilities. For $2^{2N}(2/16)^N=(1/2)^N$, a number infinitesimal compared to unity, whereas the number of arrangements must come out an integer. As a matter of fact, there are just two arrangements consistent with energy zero: one in which all dipoles point along $+c$, the other in which all are along $-c$.)

Next we find by the same method the number of arrangements with energy ϵ. In this case there is one dipole oriented perpendicular to $\pm c$, the others being parallel to this direction. There are, first, N possible dipoles any one of

[6] L. Onsager, talk presented at Conference on Dielectrics, New York Academy of Sciences, May, 1939; private communication to Dr. C. C. Stephenson.

which could be perpendicular to c. The particular dipole chosen has four chances out of its sixteen of being perpendicular to c. Each of the remaining dipoles has two chances out of sixteen of being parallel to $\pm c$. Hence we should expect $2^{2N}N(2/16)^{N-1}(4/16)$ arrangements of the hydrogens leading to energy ϵ.

(We note that this result is also certainly wrong. Like the previous one, it indicates a number of arrangements very much less than unity. Actually, there are no possible arrangements of the hydrogens at all which lead to energy ϵ. It is impossible to shift the orientation of a single dipole by itself, for this would amount to shifting the hydrogens about that phosphate group from the directions pointing to certain neighbors, to directions pointing to other neighbors. But that would leave two bonds without any hydrogens at all, and would place two hydrogens on each of two other bonds, increasing the energy so much that we are leaving such states out of account. The only way to shift the dipole of one group is to make simultaneous shifts in a whole chain of phosphate groups. It is not hard to show, from the geometry of the crystal, that such a chain cannot close on itself, but must extend from one face of the crystal to another. Thus it ordinarily will contain a number of dipoles of the order of magnitude of $N^{\frac{1}{3}}$, so that the lowest energy, greater than zero, for which there will be real arrangements of the dipoles, will be of the order of magnitude of $N^{\frac{1}{3}}\epsilon$.)

Next we find the number of arrangements with energy 2ϵ. There are $N(N-1)/2$ ways of picking out two dipoles to be perpendicular to the $\pm c$ axis. Thus, by an extension of the argument used above, we conclude that there are $2^{2N}N(N-1)/2(2/16)^{N-2}(4/16)^2$ such arrangements. And it is now easy to generalize: the number of arrangements with energy $j\epsilon$ would be expected to be

$$F(j)=2^{2N}\frac{N(N-1)\cdots(N-j+1)}{j!}(2/16)^{N-j}(4/16)^j$$
$$=2^{2N}\frac{N!}{j!(N-j)!}(2/16)^{N-j}(4/16)^j$$
$$=(1/2)^N(2)^j\frac{N!}{j!(N-j)!}. \qquad (2.1)$$

(In preceding paragraphs we have thrown doubt on the correctness of Eq. (2.1) for small values of j. As a matter of fact, the correct function $F(j)$ is an integer, which varies erratically from one value of j to the next, as we have seen by examples. For small values of j, many j values have $F(j)=0$, and a few have values different from zero. No formula like (2.1), indicating that $F(j)$ should be a continuous function of j, can be correct. If there were a formula which was correct on the average, it would have to give $F(j)$ small compared to unity for small j, since most j values have $F(j)=0$. The writer does not know whether Eq. (2.1) is even correct on the average for small j. For large j values, however, the function (2.1) becomes large compared to unity. It seems very plausible that in this range it approaches the correct number of arrangements, which is also large compared to unity. This will be made more plausible by the method of derivation used in the next section, which does not use the assumption that different phosphate groups are statistically independent of each other, as the present derivation does.)

If we accept Eq. (2.1), we can compute the partition function Z. To do this, we need merely multiply the number of states with energy $j\epsilon$, by the factor $\exp(-j\epsilon/kT)$, and sum over j, from zero to N. Thus we have

$$Z=\sum_{j=0}^{N}(1/2)^N(2\exp(-\epsilon/kT))^j\frac{N!}{j!(N-j)!}. \qquad (2.2)$$

Noting that the coefficients in the sum (2.2) are the binomial coefficients, we can carry out the summation at once, obtaining

$$Z=(1/2)^N(1+2\exp(-\epsilon/kT))^N \qquad (2.3)$$
$$=(1/2+\exp(-\epsilon/kT))^N.$$

From the partition function we can at once compute the Holmholtz free energy A, using the relation $A=-kT\ln Z$, from which

$$A=-NkT\ln(1/2+\exp(-\epsilon/kT)). \qquad (2.4)$$

The entropy S can be found by the relation $S=-(\partial A/\partial T)$, from which

$$S=Nk\ln(1/2+\exp(-\epsilon/kT))$$
$$+\frac{N\epsilon}{T}\frac{1}{\frac{1}{2}\exp(\epsilon/kT)+1}. \qquad (2.5)$$

From Eq. (2.5) we can at once get the specific heat by differentiation, in the usual way.

From Eqs. (2.3), (2.4), and (2.5) there is no indication of a discontinuous change of phase. We see at once that this is on account of the errors, which we have already discussed, in the enumeration of Eq. (2.1). For the low temperature phase should be expected to result from the very lowest energy levels, and we have failed completely in Eq. (2.1) to give the correct number of such levels, so that our partition function is incorrect at low temperatures. Really the leading term in the partition function (2.2) should be 2, the number of configurations with energy zero, so that at sufficiently low temperatures, where all Boltzmann factors $\exp(-j\epsilon/kT)$ become negligibly small, Z should approach 2, instead of approaching the infinitesimal value $(1/2)^N$ as (2.3) does. We have concluded in a previous paragraph that when the expression $F(j)$ of Eq. (2.1) is large compared with unity, it may very possibly be correct, but that when it is small compared with unity, it is certainly wrong. Let us then form the expression

$$Z'=2+Z, \tag{2.6}$$

where Z is as given in Eq. (2.3). The expression (2.6) has the correct behavior both at high and at low temperatures. It is not hard to show that when the expression $F(j)$ of (2.1) is small compared to unity, the same is true of Z of Eq. (2.3), and when $F(j)$ is large, Z is also large. Thus in the range where $F(j)$ is small compared to unity, Z' becomes equal to 2, as it should, while if $F(j)$ is large compared to unity, 2 can be neglected in comparison with Z, and again Z' takes on the correct value. But now if we use Eq. (2.6) for Z', and take $A=-kT\ln Z'$, we have the following situation: At temperatures for which $Z\ll 2$, we have approximately $A=-kT\ln 2$, which is very close to zero (since it does not contain a factor N, as does the value (2.4)). On the other hand, at temperatures for which $Z\gg 2$, the value of A is given by Eq. (2.4), to a very good approximation. The transition from one of these regions to the other comes when $Z=2$, or when $(1/2+\exp(-\epsilon/kT))=2^{1/N}=1$ approximately. In other words, for the transition temperature Θ we have

$$1/2+\exp(-\epsilon/k\Theta)=1,$$
$$\epsilon=k\Theta\ln 2. \tag{2.7}$$

This temperature Θ is the Curie temperature; It is the same one predicted by the theory of Onsager.[6] We can easily investigate the range of temperature in which the free energy A changes over from the value zero for the low temperature phase, to the value (2.4) for the high temperature phase, and can show that this range is entirely negligible.[7] Thus we can with perfect justification say that there are two phases, one with the free energy zero, one with the value (2.4); that the transition comes at the temperature where the free energy curves cross; and that the phase with the lower free energy is the stable phase.

There is an interesting fact to be noted regarding this transition. Below the Curie point Θ, the value of the function A of (2.4) is positive, for the Curie point is determined by the condition that it be zero. That is, the function Z of (2.3) is less than unity. The region below the Curie point, in other words, is the region where our method of enumeration is incorrect, since by its method of definition the correct Z can never be less than unity. Our method of calculation, in other words, gives no information about the hypothetical unstable state below the Curie point. In fact, if we could carry out the calculations correctly, we see that there would be no such unstable state; the correct value of Z would be much more like the Z' of Eq. (2.6), which automatically would bring about a sharp change of properties at the temperature Θ.

Above the Curie point, our calculation should most likely be correct. It is interesting then to work out the properties to be expected above this point. In Eq. (2.5) we see the entropy, and this should hold above Θ, though below this temperature, since A is zero, the entropy should also be zero. (It is of course to be noted that we are neglecting vibrational entropy, and specific heat.) Differentiating (2.5), we can get the contribution to the specific heat to be expected above Θ; this proves to be an amount of fairly small magnitude, distributed over a wide range of temperature. It is very different from the

[7] J. C. Slater, *Introduction to Chemical Physics* (McGraw-Hill Book Co. Inc., 1939), pp. 265–269.

narrow peak just below the Curie point which is observed, and it is almost impossible to say experimentally whether it corresponds to observation or not, since we do not know an accurate theoretical value for the remaining specific heat, resulting from vibrations. In addition to the specific heat, however, it is interesting to compute the increase of entropy on transition, and from it the latent heat. Putting in $T=\Theta$, in Eq. (2.5), we find

$$S=Nk \ln 2/2 \tag{2.8}$$

immediately above the Curie point. This gives an entropy increase of about 0.69 calorie per mole per degree at the transition. At high temperatures, the expression (2.5) approaches $Nk \ln (3/2)$, the value found by Pauling for ice. This is as it should be, for at high temperature it makes no difference whether all levels have the same energy, as in ice, or different energies as in potassium phosphate. This value is approximately 0.805 calorie per mole per degree. In other words, we should expect a sudden increase of entropy of 0.69 calorie at the transition, with an additional 0.115 calorie distributed over a wide range of temperature (hundreds of degrees). Experimentally,[8] the increase at the Curie point (distributed over several degrees) is about 0.8 calorie, rather more than we should expect. We have no explanation to suggest for this discrepancy. As we have mentioned before, it is impossible to check up the distributed entropy increase, or small specific heat anomaly above the Curie point, on account of the lack of exact knowledge of the vibrational specific heat.

3. More Accurate Theory

The theory worked out in the last section inevitably led to a phase change of the first order, with a latent heat. But we remember that we have considered really only two states, a completely ordered one (the lowest energy level), and a completely disordered one (the other levels, grouped together). Even in the case of the Weiss theory, if we considered only two such states, we should find a phase change of the first order. To get the phase change of the second order, in that case, we must consider the intermediate states, with intermediate degrees of order, and dipole moments between zero and the maximum value. We shall now do the same thing here. We shall enumerate the states of a given dipole moment, find the free energy, and see if the system really changes discontinuously from the state of maximum dipole moment to that of no dipole moment, or whether it changes gradually over a range of temperatures, with gradual decrease of dipole moment, as in magnetism. We shall find that here the change is discontinuous, even when we consider intermediate states. Thus we do not really improve our theory of the transition. We do improve the validity of our arguments, however. And at the same time, by considering the free energy of states of various dipole muments, we lay the foundations for a calculation of dielectric susceptibility above the Curie point,

Let there be, as before, N phosphate groups, each with two hydrogens. These hydrogens form a dipole, which as we have said before can point along the $+c$ direction, perpendicular to the c direction (in either of four orientations), and along $-c$. Let us consider an arrangement of the hydrogens in which N_+ dipoles point along $+c$, N_0 along a perpendicular direction, and N_- along $-c$. Plainly

$$N_+ + N_0 + N_- = N. \tag{3.1}$$

Our problem is now to find the number $F(N_+, N_0, N_-)$ of arrangements of the hydrogens, leading to these values of the N's. There is obviously a close connection between this F, and the $F(j)$ of Eq. (2.1); we shall discuss the relation later. The quantity j of Section 2 is easily seen to be the same as the N_0 introduced here. The advantage of the present treatment is that it allows us to find how many arrangements there are with a given dipole moment, which depends on $N_+ - N_-$.

We shall now assume that our function F can be approximated in the form

$$F(N_+, N_0, N_-) = a^{N} b^{N_0} c^{N_-}, \tag{3.2}$$

where a, b, c are functions of the ratios of N_+, N_0, N_- to N, which we shall abbreviate as follows:

$$\nu_+ = N_+/N, \quad \nu_0 = N_0/N, \quad \nu_- = N_-/N, \tag{3.3}$$

[8] C. C. Stephenson, unpublished measurements.

so that

$$\nu_+ + \nu_0 + \nu_- = 1. \tag{3.4}$$

Thus, we assume $a = a(\nu_+, \nu_0, \nu_-)$, etc. The justification for the assumption (3.2) is that almost certainly F can be set up as a function involving factorials, as was Eq. (2.1), which can be approximated in the form (3.2) by using Stirling's theorem. We shall later show that (2.1) can be written in the form (3.2). Our problem, then, is to find the functions a, b, and c. We shall do this by a method of building up the crystal molecule by molecule. We shall assume that at a given step in the process of building up, it contains a certain number of molecules, and that we are then adding another molecule to the surface. We ask, essentially, how does the number of arrangements increase on account of the addition of this new molecule, considering the various ways in which the hydrogens of this molecule can be oriented? This gives us a connection between the function F for a certain number of molecules, and the corresponding function for a number greater by unity, and this allows us to evaluate the functions a, b, and c in a simple and unique way.

In Fig. 2, we symbolize two successive layers of phosphate groups in the crystal, looking down along the c axis. The dotted lines symbolize hydrogen bonds, and we recall that each refers to a hydrogen that can be near one or the other of the two oxygens which the bond joins, subject to the conditions that two and only two hydrogens are attached to each phosphate group. Each bond, as we see, connects the upper corner of a phosphate group in the lower layer, and the lower corner of a group in the upper layer. In addition to these two completed layers, we show a group, symbolized by α, in the layer above these. This group α is tied by hydrogen bonds to two groups in the layer below, symbolized by β and γ, the corresponding hydrogen bonds being denoted by (1) and (2). Now we note that the orientations which the dipole of group α can take on are almost uniquely determined by the dipoles already present on β and γ. Let us refer to a hydrogen as $+$ if it is near the upper corner of the phosphate group in the layer below, and as $-$ if it is near the lower corner of the phosphate group in the layer above; thus if

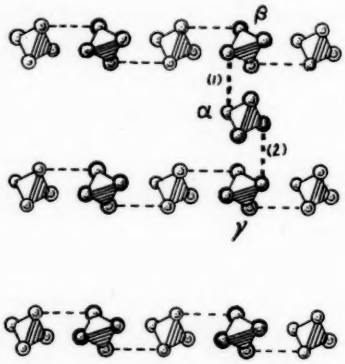


FIG. 2. Two layers of tetrahedra, looking down along the c axis. The upper layer is shaded more heavily. The tetrahedron α is in a layer above the two which are shown. Dotted lines represent hydrogen bonds.

the two hydrogens attached to a phosphate group are both $+$, the dipole moment points along $+c$; if one is $+$, one $-$, it is perpendicular to c; while if both are $-$, it points along $-c$. Then we see that if hydrogens (1) and (2) are both $+$, they are attached to the upper corners of β and γ, respectively, leaving the two lower corners of α without hydrogens. Thus the two upper corners of α must have hydrogens, so that α must have its dipole pointing along $+c$. If hydrogen (1) is $+$, (2) $-$, or vice versa, group α must have one hydrogen at one upper corner (since it has one at a lower corner), and this hydrogen can be at either corner, giving two possibilities, leading to the four possible orientations of the dipole α perpendicular to c. Finally, if hydrogens (1) and (2) are both $-$, the dipole of α points along $-c$.

We next ask, what are the probabilities of these various arrangements? First what are the probabilities that the hydrogen (1) be $+$ or $-$? In other words, in what fraction of all the arrangements of dipoles is this particular hydrogen $+$, and in what fraction is it $-$? At this point, we shall definitely limit the validity of our arguments to cases where there are many arrangements consistent with a given set of N's; that is to say, to cases where F is large compared with unity. This is essentially the same limitation to which we were led in the more elementary treatment of Section 2; there we found that our calculation was certainly wrong when F was comparable with unity, or small compared to unity. But if F is large, and there are many arrangements of the dipoles, then there will be certain normal properties which a great majority of the arrangements will possess. One such property which we can certainly expect is the

following: If a fraction ν_+ of all dipoles points along the $+c$ axis, then the fraction of all arrangements in which a particular dipole, as β, will point along this direction will be ν_+. This may be expected to be independent of which dipole β is chosen; that is, there is a homogeneity throughout the crystal. We shall shortly apply a condition stating that each successive layer of atoms, as we add it to the crystal, shall have the same properties as the crystal as a whole, thus securing this homogeneity, and assuring us of the fact that the group β does not have special properties on account of being on the surface of the crystal. Now we note that all the arrangements of dipoles in which the dipole β points along $+c$, and half those in which it is perpendicular to c, correspond to the hydrogen (1) being $+$. Thus we expect the fraction of all arrangements in which (1) is $+$, or the probability that (1) be $+$, to be equal to $(\nu_+ + \frac{1}{2}\nu_0)$. If we let this probability be p, then

$$p = \nu_+ + \tfrac{1}{2}\nu_0, \quad 1-p = \nu_- + \tfrac{1}{2}\nu_0. \tag{3.5}$$

Here p is the probability that hydrogen (1) be $+$, $1-p$ is the probability that it be $-$.

The probabilities that hydrogen (2) be $+$ or $-$, respectively, are surely the same as those given in (3.5). We next ask, what is the probability that simultaneously (1) and (2) should be $+$? In other words, are these two events independent, so that the combined probability is simply the product of the probabilities of the separate events, or p^2? There obviously are some cases in the crystal where two hydrogen bonds cannot be treated as independent of each other. Thus, for instance, if both bonds are attached to the same phosphate group, and we know the dipole moment of that group, there is certainly a very strong relation between the two bonds. But in the present case, before the attachment of group α (and that is the case we are considering), hydrogens (1) and (2) are attached to groups β and γ, respectively, groups which are in different rows in the crystal, and which are not tied together in any direct way whatever by hydrogen bonds. Thus, although they are close together geometrically, they would be expected to be almost entirely independent of each other in the matter of their hydrogens. We should, therefore, be justified in making the following statements:

The probability that hydrogens (1) and (2) are both $+$ is p^2.

The probability that hydrogen (1) is $+$, and (2) is $-$, is $p(1-p)$; similarly the probability that (1) is $-$, (2) $+$, is $p(1-p)$.

The probability that (1) and (2) are both $-$ is $(1-p)^2$. We note, as we expect, that the sum of these probabilities is $p^2+2p(1-p)+(1-p)^2=1$.

Now we can return to our function $F(N_+, N_0, N_-)$, the number of arrangements of hydrogens when there are N_+, N_0, N_- dipoles pointing along $+c$, the perpendicular directions, and $-c$, respectively. Assume that the whole number N of dipoles includes the dipole α of Fig. 2. We shall now find the relation between this whole number of arrangements, and the smaller number which we should have if α were absent. We shall write F as the sum of three numbers: the number of arrangements in which α points along $+c$, the number in which it is perpendicular to c, and the number in which it is along $-c$. First, how many arrangements are there with α along $+c$? These are just the arrangements in which hydrogens (1) and (2) are both $+$, and in which, out of the $N-1$ dipoles of the incomplete crystal, N_+-1 point along $+c$, N_0 perpendicular, and N_- along $-c$. In all, there are $F(N_+-1, N_0, N_-)$ arrangements of the incomplete crystal, of which the fraction p^2 corresponds to having hydrogens (1) and (2) both $+$. Thus the desired number is $p^2F(N_+-1, N_0, N_-)$. Next, how many arrangements are there with α perpendicular to c? There are $2p(1-p)F(N_+, N_0-1, N_-)$ arrangements in which one of the hydrogens is $+$, the other is $-$, and each of these corresponds to two possible arrangements of the hydrogens of dipole α. Thus the desired number of arrangements is $4p(1-p)F(N_+, N_0-1, N_-)$. Finally the number of arrangements in which the dipole α points along $-c$ is $(1-p)^2F(N_+, N_0, N_--1)$. And now, if we write the total number of arrangements as the sum of these, we have

$$F(N_+, N_0, N_-) = p^2F(N_+-1, N_0, N_-) + 4p(1-p)F(N_+, N_0-1, N_-) + (1-p)^2F(N_+, N_0, N_--1). \tag{3.5}$$

Using Eq. (3.2) for F, this can be rewritten

$$(a^{\nu_+}b^{\nu_0}c^{\nu_-})^N=\frac{p^2}{a}(a^{\nu_+}b^{\nu_0}c^{\nu_-})^N+\frac{4p(1-p)}{b}(a^{\nu_+}b^{\nu_0}c^{\nu_-})^N+\frac{(1-p)^2}{c}(a^{\nu_+}b^{\nu_0}c^{\nu_-})^N, \tag{3.6}$$

from which at once

$$1=\frac{p^2}{a}+\frac{4p(1-p)}{b}+\frac{(1-p)^2}{c}. \tag{3.7}$$

And the fraction of all arrangements in which the dipole α is along $+c$ is p^2/a; the fraction in which it is perpendicular to the c axis is $4p(1-p)/b$; and the fraction in which it is along $-c$ is $(1-p)^2/c$.

Next we can apply the condition which requires our crystal to be homogeneous in its properties: The fraction of all arrangements in which the dipole α points in a given direction must be the same as the fraction of arrangements in which any other dipole of the crystal points in that direction. But we already have assumed that the fraction of all arrangements in which a dipole points along $+c$ is ν_+; the fraction in which it is perpendicular to c is ν_0; and the fraction in which it is along $-c$ is ν_-. Hence we have the relations

$$p^2/a=\nu_+, \quad 4p(1-p)/b=\nu_0, \quad (1-p)^2/c=\nu_-. \tag{3.8}$$

From Eq. (3.8) we can immediately evaluate the unknown functions a, b, and c; using Eq. (3.5), these become

$$\begin{aligned} a&=p^2/\nu_+=\frac{(\nu_++\frac{1}{2}\nu_0)^2}{\nu_+},\\ b&=\frac{4p(1-p)}{\nu_0}=\frac{4(\nu_++\frac{1}{2}\nu_0)(\nu_-+\frac{1}{2}\nu_0)}{\nu_0},\\ c&=\frac{(1-p)^2}{\nu_-}=\frac{(\nu_-+\frac{1}{2}\nu_0)^2}{\nu_-}. \end{aligned} \tag{3.9}$$

And finally, using Eq. (3.2), we evaluate the function F, obtaining

$$F(N_+, N_0, N_-)=\left\{\left[\frac{(\nu_++\frac{1}{2}\nu_0)^2}{\nu_+}\right]^{\nu_+}\left[\frac{4(\nu_++\frac{1}{2}\nu_0)(\nu_-+\frac{1}{2}\nu_0)}{\nu_0}\right]^{\nu_0}\left[\frac{(\nu_-+\frac{1}{2}\nu_0)^2}{\nu_-}\right]^{\nu_-}\right\}^N. \tag{3.10}$$

With the evaluation of the function F in Eq. (3.10) we have solved the most difficult part of our problem. The rest is relatively straightforward statistical mechanics and thermodynamics. We have the choice of two methods. We could compute the partition function Z, as in Section 2, and find the Helmholtz free energy from that. Or we could take the entirely equivalent, but slightly more comprehensible, method of finding the entropy from the number of arrangements in Eq. (3.10), and find the Helmholtz free energy from it by the formula $A=U-TS$, getting A in a form involving the N's as parameters. Then we can determine the N's so as to make A a minimum, since we know that in equilibrium A takes on the lowest possible value. We shall adopt this second method.

We first find the free energy of a state with arbitrary N's. Its entropy is given by Boltzmann's formula in terms of the number F of arrangements consistent with these N's:

$$S=k \ln F(N_+, N_0, N_-). \tag{3.11}$$

The internal energy, if we neglect the thermal vibrations, comes wholly from the N_0 phosphate groups which have their moments perpendicular to the c direction. Each such group has an energy ϵ, leading to an internal energy $U=N_0\epsilon$ for the whole crystal. This is in the absence of an external electric field. If, however, there is a field, along the $+c$ axis, of magnitude E, there will be an additional energy, equal to the negative of the product of E and the electric moment of the crystal along $+c$.

If μ is the dipole moment of a group along $+c$, this total dipole moment is $(N_+-N_-)\mu$. Thus the additional electric energy is $-(N_+-N_-)\mu E$. Hence the total internal energy is

$$U=N_0\epsilon-(N_+-N_-)\mu E. \tag{3.12}$$

In terms of the U and S given in Eqs. (3.11) and (3.12), the Helmholtz free energy A is then given by

$$A=U-TS. \tag{3.13}$$

In Eqs. (3.11), (3.12), and (3.13), we have expressed A as a function of the three parameters N_+, N_0, N_-. Of these three, only two are independent, since the three must add to N. In place of these quantities, we shall use the parameters

$$x=\nu_+-\nu_-, \tag{3.14}$$

and ν_0, as parameters; the quantity x is convenient since it is a measure of the net dipole moment, being zero for the unpolarized crystal, unity for the completely polarized crystal, in which all dipole moments point along $+c$, and -1 for the case where all dipoles point along $-c$. In thermal equilibrium, these two parameters will be determined by the condition that A be a minimum with respect to variation of either one. We shall first apply this condition to ν_0, differentiating (3.13) with respect to ν_0, and setting the derivative equal to zero. This will give us a value of ν_0, which we can insert in A, resulting in a function of x only, which we shall then discuss explicitly, rather than carrying out an additional differentiation with respect to x.

It is a matter of simple mathematics to carry out the differentiation with respect to ν_0, and we shall not reproduce the mathematical steps. Using Eqs. (3.3), (3.4), (3.10), (3.11), (3.12), and (3.13), we can express A in terms of x and ν_0. Then we differentiate with respect to ν_0, keeping x constant, (as well, of course, as keeping T and E constant), and set the derivative equal to zero. This fortunately gives a rather simple equation, a quadratic in ν_0, which can be solved at once. It is a simple matter to find which of the two solutions of the quadratic we wish. The result then proves to be

$$\nu_0=\frac{4-2[x^2(4-\exp(2\epsilon/kT))+\exp(2\epsilon/kT)]^{\frac{1}{2}}}{4-\exp(2\epsilon/kT)}. \tag{3.15}$$

If we use this equilibrium value of ν_0, substituting in Eq. (3.13), we have a value of A in which x is the only parameter. The resulting function is obviously rather complicated mathematically, and for that reason it is worth while to make various power series developments, which allow us to understand its properties better. It is rather a tedious mathematical problem to make these series developments, but again it is perfectly straightforward, and for that reason we give only the results, not the intermediate steps. In the first place, as in Eq. (2.7), it turns out that the transition temperature Θ is given by the relation $\epsilon=k\Theta \ln 2$, or $\exp(\epsilon/k\Theta)=2$. This suggests that, in place of the temperature, it is convenient to use the variable

$$y=2-\exp(\epsilon/kT), \tag{3.16}$$

which is zero at the transition point, and increases as the temperature increases, going to 1 at infinite temperature. We then carry out two related, but somewhat different, expansions: First we expand the function A/NkT (aside from the term involving the electric field) in power series in y, expressing each coefficient exactly as a function of x; secondly, we expand A/NkT in power series in x, expressing each coefficient exactly as a function of y. These expansions are

$$A/NkT=-x\mu E/kT-y(1-x^2)/4+y^2(-\tfrac{3}{32}+\tfrac{1}{8}x^2-\tfrac{1}{32}x^4)\cdots, \tag{3.17}$$

$$A/NkT=-\ln(\tfrac{1}{2}+\exp(-\epsilon/kT))-x\mu E/kT+x^2(\exp(-\epsilon/kT)-\tfrac{1}{2})+\cdots. \tag{3.18}$$

Plainly we should expect that if we expand the coefficients of the terms of (3.18) in power series in y, the first few terms of both series should agree; when this check is made, it is found that they do.

First we consider the free energy in the absence of an external field, the case $E=0$. This is plotted, for various temperatures, as a function of x, in Fig. 3. The first thing we notice is that for $y=0$, or $T=\Theta$, the curve is horizontal, corresponding to $A=0$. This is most easily seen from the expansion (3.17). For $T>\Theta$, the curve has a minimum at $x=0$, and for $T<\Theta$, there is a maximum for $x=0$, corresponding to a positive value of A. For $x=1$, the value of A is zero in every case. This can be seen easily directly from fundamentals. The case $x=1$ is that in which all dipoles are along the $+c$ direction. There are then none perpendicular to c, so that the internal energy is zero. Furthermore, there is just one such arrangement, so that we should expect that the factor F would be unity, and S would be zero, resulting in zero for the free energy. As a matter of fact, our function F of Eq. (3.10) does come out unity in this case, as it should; showing that the method of enumeration used in the present section is more reliable than the elementary one of Section 2, since that gave the incorrect answer in the case of this lowest energy state. Now we can consider the condition that for equilibrium A must take on the lowest possible value. Plainly for $T<\Theta$, the lowest states are for $x=1$ or -1, corresponding to complete polarization of the crystal. For $T>\Theta$, on the other hand, the state of $x=0$ has the lowest free energy. Thus at this temperature there will be a sudden transition from the polarized to the unpolarized state. It is interesting to consider the contrast between this case and that of the Weiss theory. In that theory, the corresponding curves for free energy take the form shown in Fig. 4. Above the Curie point the minimum comes for $x=0$, but below that point, instead of suddenly shifting to $x=\pm 1$, it gradually shifts, not reaching complete polarization until the temperature reaches the absolute zero. For that reason, the Weiss theory predicts a phase change of the second order, with a gradual transition, whereas the present theory leads to a sudden transition. We discuss the relation of this result to experiment in a later paragraph.

Above the transition point, the stable state is that for $x=0$. In that case, we can find the free energy from the first term of Eq. (3.18). We now notice that this is the same that was found by the elementary theory of Section 2, in Eq. (2.4). Thus the discussion of entropy and specific heat which was given there holds for this more correct theory, and we see that, though the elementary theory seemed unsatisfactory because it appeared to neglect the interactions between phosphate groups, still it really led to essentially the right result. We can see why this happened, by comparing our function $F(N_+, N_0, N_-)$ of (3.10) with the function $F(j)$ of Eq. (2.1). To get the unpolarized state in (3.10) we should set $x=0$, or $\nu_+=\nu_-$, in (3.10). Then $\nu_+ + \frac{1}{2}\nu_0 = \nu_- + \frac{1}{2}\nu_0 = \frac{1}{2}$, so that F becomes

$$\left\{ \frac{(1/4)^{1-\nu_0}}{\nu_+^{\nu_+}\nu_0^{\nu_0}\nu_-^{\nu_-}} \right\}^N .$$

But taking the value $F(j)$ of Eq. (2.1), setting $j=N_0$, and using Stirling's formula for the factorials, we get a value which can be shown simply to have the same value. This verifies the statement made early in this section about the relations between the two functions F, and the fact that $F(j)$ of (2.1) can be written in the form (3.2). The method of enumeration in the preceding section, in other words, gives correct results for $x=0$, but it cannot be adapted to give results for other values of x.

In the presence of an external field, the equilibrium will be shifted, so that above the transition point the stable state will no longer come for $x=0$. To find the susceptibility for small external fields, we can most easily use Eq. (3.18), and carry out our differentiation with respect to x, setting the derivative equal to zero for equilibrium. When we do this, we find at once

$$x = \frac{\mu E}{2kT(\exp(-\epsilon/kT) - 1/2)}. \tag{3.19}$$

This gives a polarization which becomes infinite at the transition temperature, as in Curie's law. The denominator of (3.19) can be expanded in powers of $T-\Theta$, and the first term is $k \ln 2(T-\Theta)$. Using this value, the polarization $N\mu x$ becomes

$$N\mu x = \frac{N\mu^2}{k \ln 2} \frac{E}{T-\Theta}. \tag{3.20}$$

This is to be compared with the value given by

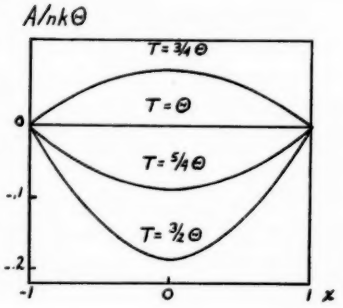


FIG. 3. Free energy as function of dipole moment, at different temperatures, for KH_2PO_4.

the Weiss theory,

$$N\mu x=\frac{N\mu^2}{3k}\frac{E}{T-\Theta}. \qquad (3.21)$$

In other words, the polarization in the present case is $3/\ln 2=4.33$ times as great as it would be according to the Weiss theory. It is interesting to ask why there is this large difference. It is easy to see that if the free energy curves of Fig. 3 have very deep, sharp minima, the polarizability will be small, and vice versa. For with a deep minimum the effect of superposing a linear curve (coming from the external field) on the curve of Fig. 3 will only shift the minimum slightly. Conversely with a shallow minimum a field will have a large effect. But now if there is a large entropy difference between the unpolarized and the polarized states, the free energy of the unpolarized state will decrease much more rapidly with temperature than for the polarized state, so that the minimum will be sharp at a given temperature above the Curie point (for we remember that $S=-(\partial A/\partial T)$). Conversely, if there is a small entropy difference between states, the minimum will be shallow. The Weiss case is that in which there is no coupling between dipoles, so that all arrangements of the dipoles are allowed, and the entropy of the unpolarized state is very large. In our present case, however, the number of arrangements is greatly restricted by the coupling between dipoles, which forbids arrangements in which two hydrogens are located along the same bond. In other words, the entropy of the unpolarized state is not nearly as large as in the Weiss case, and as we have just seen, this results in a larger polarizability. We shall see in Section 5 that this large polarizability is in agreement with experiment.

4. Discussion of the Transition, and Comparison with Experiment

Let us first review the results of the preceding sections. We have found the free energy of our crystal as a function of x, which measures the dipole moment as a fraction of the maximum possible, or saturation moment, and of the temperature, in the absence of an external field. Remembering that the stable state is that of minimum free energy, we find that below a transition temperature Θ, the completely polarized state, corresponding to $x=\pm 1$, is the stable state, while above this temperature the unpolarized state is the stable one. The entropy increase in the transition should be $R(\ln 2)/2$. Above the transition, spread through a wide temperature range, there should be an additional entropy increase, enough to raise the entropy from $R(\ln 2)/2$ (which is 0.69 cal.) to $R \ln 3/2$ (which is 0.805 cal.). Above the transition point, in the presence of an external electric field, there will be a polarization proportional to the field, given by Eq. (3.20), similar to the value of the Weiss theory, but with a coefficient 4.33 times as great.

Experimentally, the transition is not sudden, but takes place through a considerable temperature range. It is much more concentrated in a few degrees, however, than the Weiss theory would give, so that it seems possible that it is merely a phase change of the first order, distorted in some way. This is made even more plausible

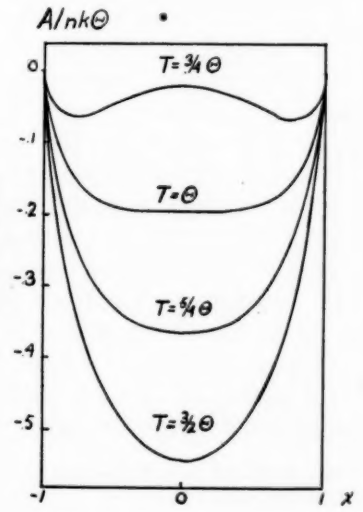


FIG. 4. Free energy as function of dipole moment, at different temperatures, according to Weiss theory.

by the fact that the transitions in the crystals KH_2AsO_4 and $(NH_4)H_2PO_4$,[9] which have the same structure and the same sort of transition, are much sharper than in KH_2PO_4. We ask, then, for possible causes of the observed gradual nature of the transition. In the first place, is it possible that our theory is inadequate, and that it really should predict a phase change of the second order? This, of course, is possible. We have noted that our calculation of the partition function and free energy of the high temperature phase gets more and more doubtful as we approach the transition point, for the number of possible arrangements of dipoles, which we use in finding the entropy, decreases toward unity and below, and in our calculations we have used methods which are only correct when the number of arrangements is large. Thus it is always possible that we have made errors which, if corrected, would change the sudden transition into a gradual one. The writer feels, however, that this is not likely to be the case, and that even a more correct calculation would still show the sudden transition in the same way. There is one additional way in which the theory might possibly be in error. We have explicitly neglected the possibility that there may be two hydrogens, or no hydrogens, on the bond joining two oxygens; we have allowed only the case where there was just one hydrogen on each bond. Perhaps we should have considered these other possibilities. They would certainly have increased the entropy of the unpolarized state (since they provide other possible arrangements), and thus they might help in some measure to bring the theoretical entropy of transition from the value 0.69 calorie, which we have found, to the experimental value of 0.80. And without calculation we cannot be sure they would not have the effect of rounding off the transition. Here again the writer is skeptical, however. These possibilities, as has been mentioned in Section 1, would correspond to the formation of ions, a type of process which is known to be very rare in the similar case of ice, and is probably rare here too, so that though the effect undoubtedly exists, it is very doubtful whether it could be big enough to account for the observations.

[9] See reference 1 for KH_2AsO_4; also C. C. Stephenson, unpublished work.

It seems more plausible to ascribe the gradual nature of the transition to physical causes of some sort, which we have not considered in our theory. Of course, impurities would have just this effect, but there is no reason for suspecting impurities in the samples used in the experiments. But the writer is inclined to think that the most plausible explanation is the following. The crystal is known experimentally[10] to have an abnormally high piezoelectric effect, with correspondingly large inverse piezoelectric effect, both increasing greatly as the temperature decreases to the transition temperature. That is, a given electric field will distort the crystal geometrically, by an amount which becomes quite considerable as the transition temperature is approached. And if we look into the term in the free energy giving the coupling between electric and elastic energy (a term which we have not included in our calculation), we see that we expect a spontaneous distortion of the crystal below the transition. The symmetry properties tell us the sort of distortion this will be: It will be a shear, the squares of Fig. 2 being distorted into diamond shaped figures, the shear being in opposite directions for polarizations along the $+c$ and $-c$ axes. But we cannot assume that a real crystal, in cooling through the transition point, will take on such a shear. The reason is that the real crystal below the transition point undoubtedly consists of domains, some polarized in one direction and some in the other. This is shown, for instance, by the fact that the crystal shows electric hysteresis, like the magnetic hysteresis of ferromagnetic bodies, and that the normal state of the crystal below the transition temperature is unpolarized. If there are domains, some would tend to be sheared in one direction, some in the other, which they could not do without breaking the crystal. Since the indicated shear is small, what presumably happens is that each domain is constrained by its neighbors to very nearly the undistorted shape, so that it is under a considerable stress in the crystal below the transition point. This is made very likely by the observation[11] that when a crystal below its transition point is run rapidly through a hysteresis curve, it is very likely to shatter, as if it

[10] W. Lüdy, Zeits. f. Physik **113**, 302 (1939).
[11] G. Busch, Helv. Phys. Acta **11**, 269 (1938).

were under internal stress. But such stresses would alter the transition point; as a matter of fact, they would depress the transition point. Perhaps the more accurate way to state the effect is as follows: Our theory as we have worked it out in preceding sections is correct for the unpolarized high temperature phase of the crystal, and would be correct for the polarized low temperature phase if that did not distort. The low temperature phase actually does distort, however. That means that the stable, distorted state must have a lower free energy than the unstable, undistorted state. Hence the real free energy of the high temperature phase is as we have found it, but the real free energy of the low temperature phase is lower than we have found it. In Fig. 3, the correct point for $x=1$, in other words, should be for A slightly negative, while our value for $x=0$ is correct. It is plain, then, that the two free energies will be equal, and the transition will occur, at a somewhat higher temperature than the transition point we calculated previously. In other words, the transition point of the unstressed material is higher than that of the stressed material, or stress lowers the transition point. This effect is similar to effects in Rochelle salt, recently discussed by Mueller.[12]

Our hypothesis, then, is that the actual crystal contains domains under varying degrees of stress, simply on account of their desire to distort, which is hindered by their being fastened together, and by the fact that they do not distort enough to pass the elastic limit or break. And the stresses depress the transition points of these domains by various amounts, so as to give a continuous transition process. Unfortunately the elastic constants for KH_2PO_4 are not well enough known so that we can calculate exactly the change of transition point to be expected from reasonable stresses. Rough estimates, however, indicate that if the crystal were clamped so that it could not distort in the transition, the transition point would be lowered by a number of degrees. The effect of stresses resulting from piezoelectric deformation in the spontaneously polarized phase, then, would be to depress the transition points of domains by amounts of the right order of magnitude to account for the observed broadening of the transition.

If the mechanism we have suggested should prove to be the correct explanation of the broadening of the transition into a phase change of the second order in this case, it seems not unlikely that it might play a part in many other lambda-point transitions as well. For consider any phase change of the first order in which there is no major change of crystal structure between the phases; that is, in which there is only a change in the positions of hydrogens, or some such minor change. Even if there is no large change of structure, there is almost sure to be a change on a small scale, probably too small to be detected by x-ray measurements. With a large change in structure, the crystal normally shatters and breaks up on going through the transition, but with such a small change it may hold together, and not be strained beyond the breaking point. In this case, stresses will appear, and they will change the transition point, by amounts depending on the magnitude of the stress, so that different domains will have different transition points, and the transition will be broadened. This mechanism is not at all dependent on the special feature which we have in KH_2PO_4, the spontaneous polarization; it depends only on the existence of some very slight change in structure between the phases. It seems not impossible to the writer that a mechanism of this sort may be responsible for many of the very common rather sharp but not discontinuous lambda-point transitions, in which the change of properties is very much more sudden than is indicated by the Weiss theory, or by the closely related Bragg and Williams theory of order and disorder.

5. Susceptibility Above the Transition Point

There are good measurements of susceptibility above the transition point,[11] and we shall now compare the results of our theory with these experiments. This comparison should hardly depend on the details of the transition, the question as to whether it is sharp or not, and for that reason it should be a good test of the theory. In Eq. (3.20), we have the induced dipole moment in terms of the field, the unknown quanti-

[12] H. Mueller, Phys. Rev. **57**, 829 (1940).

ties being only the dipole moment per molecule, and the transition temperature. We can find the dipole per molecule from the saturation moment, and can take the observed transition temperature, so that we can calculate the induced dipole moment without any arbitrary constants. If we could only calculate the dielectric constant directly from this, we should be able to get an unambiguous comparison with experiment. Unfortunately, however, the situation is not quite as simple as this, on account of the polarizability of the remaining parts of the crystal, apart from the hydrogen bonds. This polarizability is not great enough to change the dielectric constant greatly by itself; but it can make a considerable change in the effective field acting on a hydrogen bond, and this in turn affects greatly the induced dipole moment of the hydrogens. We shall develop a simple theory to take account of this.

If we used Eqs. (3.19) and (3.20) directly, we should expect the susceptibility χ, which equals the polarization $N\mu x$ divided by the field E, to have the value

$$\chi_0=\frac{N\mu^2}{2kT(\exp(-\epsilon/kT)-1/2)}=\frac{N\mu^2}{2kT((1/2)^{\Theta/T}-1/2)}, \quad (5.1)$$

which can be approximated as

$$\chi_0=\frac{N\mu^2}{k\ln 2}\frac{1}{T-\Theta}. \quad (5.2)$$

We may expect two corrections to this simple theory, however. In the first place, there is an additional susceptibility χ_1, arising from the polarization of other parts of the lattice, principally the oxygens in the phosphate groups. While this will be unimportant compared to χ_0 at temperatures near the Curie point, where χ_0 is very large, it will persist to high temperatures, where it will be the principal part of the susceptibility. The experiments definitely indicate the presence of such a term. Then, in addition, we cannot properly assume that the electric field at a hydrogen bond is the same as the average field E throughout the crystal. In this connection, of course we think first of the Lorentz correction, according to which the field at a dipole should be $E+4\pi P/3$. We can hardly accept this correction as it stands, however; in Section 1 we have explicitly stated that we do not believe the Lorentz correction to be applicable to the present problem, because the field at one dipole due to all other dipoles does not point in the direction of P, but rather in a direction determined by the nearest neighbors, and in the last analysis by the orientation of the dipole we are considering, itself. When we consider the polarizability of the oxygens as well as of the hydrogen bonds, however, this point of view seems a little extreme. The polarized oxygens will produce a field, proportional to E, and independent of temperature, at each hydrogen bond. Thus the field at a hydrogen bond will not be E, but a factor times E. It is difficult to estimate this factor. It certainly will be less than if we applied a $4\pi/3$ correction, using only the oxygen polarization vector. For the $4\pi/3$ correction would give the field at a phosphate group due to all other phosphate groups. This is not what we want; we want the field at a hydrogen bond, between phosphate groups, due to all phosphate groups, and since such a position is much more like an average position in the lattice, the field there must be much more like the average field through the lattice, or E. We shall simply assume that the field at a hydrogen bond is aE, where a is a constant independent of temperature, presumably of the order of magnitude of unity. Then we have the formula

$$\chi=\chi_1+a\chi_0, \quad (5.3)$$

where χ_0 is given by (5.1) or (5.2), and where χ_1, a, are constants.

Experimentally,[11] the curve of susceptibility against temperature is of a form which can be approximated fairly well by Eq. (5.3). In the first place, it is found that the saturation polarization at low temperatures is 4.3×10^{-6} coulomb/cm^2. Assuming that this corresponds to a state where all dipoles are parallel, as our theory indicates that it should, this gives $\mu=1.25\times10^{-18}$ e.s.u. Then, using the number of molecules per cc in the crystal, Eq. (5.2) becomes $\chi_0=170/(T-\Theta)$, with a corresponding value for (5.1). Using the function (5.1) for χ_0, we then find that we get fairly satisfactory agreement

with experiment by taking the constants $\chi_1=0.60$, $a=1.80$. (Since Busch[11] presents only a graph of his results, rather than a table of values, it is hard to give an exact comparison. The agreement of our formula with his experiment is not within experimental error at the higher temperatures, but it is within three or four percent at all temperatures, which seems satisfactory for the present purpose.) Let us see if these constants seem reasonable. In the first place, if the crystal contained only the oxygens, not the hydrogen bonds, the dielectric constant would be given by $\epsilon_1=1+4\pi\chi_1=8.6$. This seems like an entirely reasonable value for such a crystal. Next we may consider the constant a. According to it, the average field at a hydrogen bond is 1.8 times the average throughout the crystal. We may compare this with the average field which there would be at a phosphate group, if the hydrogens were not present. This would be determined by a Lorentz correction in the following way: The polarization would be $\chi_1 E=0.60E$, and the added field at a phosphate group would be $4\pi/3(0.60E)=2.50E$, so that the whole field would be $3.50E$. In other words, the factor 1.80 by which we must multiply the average field to get that at a hydrogen bond is about half the factor 3.50 by which we should have to multiply it to get the field at a phosphate group, if the hydrogen bonds were absent. This checks with our impression that the hydrogen bonds were more nearly at an average position than the oxygens of the phosphate groups, and that consequently the factor would be more nearly unity for them. In other words, though we attempt no quantitative calculation, the factor a which we find seems reasonable. But it should be noted that this result comes about only because the numerical factor in our formula (3.20) for polarization is, as we have pointed out, as much as 4.33 times as great as the factor in formula (3.21), that given by the Weiss theory. If we tried to use the Weiss theory, and set up corrections to it by means of constants χ_1 and a as we have done here, our revised constant a would have to be $4.33\times1.80=7.8$, a very unreasonably large value. In other words, our statistical treatment, taking account of the interaction of dipoles, has led to a much larger polarizability than the Weiss theory, and experiment indicates this large polarizability. It is well known that values of dipole moment determined from saturation polarization below the Curie point, when used in the Weiss theory, very often do not give correct values for the susceptibility above the Curie point, in magnetic as well as dielectric cases. It may well be that the explanation is usually similar to what we have here, that the dipoles are really coupled together, and that the Weiss theory, which assumes that as far as their entropy is concerned they are independent, and which considers coupling only through the energy term involved in the Lorentz correction, is not really applicable in most cases.

The writer wishes to indicate his indebtedness to Dr. C. C. Stephenson, who pointed out the interest of the problem to him, and has generously furnished unpublished data on specific heats; and to Professor Hans Mueller, who has given valuable suggestions and advice.

The Sorting of Mixed Solvents by Ions*

GEORGE SCATCHARD
Research Laboratory of Physical Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received October 10, 1940)

The electrostatic interaction of ions with nonelectrolytes is computed for ethanol-water mixtures over the whole composition range by the Debye treatment which takes into consideration the heterogeneity of the solvent. A simple approximate relation is derived which may be extended to other solvents. The agreement with the experimental measurements is shown to be unsatisfactory. The possibility that the lack of agreement may be due to neglect of the discrete structure of the solvent is discussed qualitatively.

THE explanation of salting out of nonelectrolytes and related phenomena by the interaction of the ionic charge with the nonelectrolyte was first made by Debye and McAulay.[1] Their treatment of a nonelectrolyte mixture as a homogeneous medium with uniform dielectric constant leads to the equation for the chemical potential of an ion which is usually attributed to Born,[2]

$$E_i=\epsilon^2 z_i^2/2b_iD \qquad (1)$$

in which E_i is the electrostatic contribution to the chemical potential per ion of the ith species in a medium of dielectric constant D, ϵ is the protonic charge, z_i is the valence and b_i the radius of an ion of the ith species. Debye[3] has outlined a treatment which takes into account the heterogeneity of mixed solvents, and has calculated the effect for the limiting case of a solution infinitely dilute in both electrolyte and the nonelectrolyte of lower dielectric constant. In recent papers various kinds of deviations from Eq. (1) have been attributed to the effect of heterogeneity. It has therefore seemed worth while to calculate the effect over the whole range of nonelectrolyte composition for at least one pair of solvents in spite of the fact that no simple solution is possible in the general case. The pair selected is water-ethanol at 25°C, for many electrolytes have been studied in these mixtures, and the required properties have all been measured.

The electrostatic contribution to the free energy of an ion may be written

$$E_i=\int[c_1(\mu_1-\mu_1^0)+c_2(\mu_2-\mu_2^0)+\epsilon^2 z_i^2/8\pi Dr^4]dS, \qquad (2)$$

in which c_1 and c_2 are the concentrations of the two nonelectrolyte species in the volume element dS, at the distance r from the central ion, μ_1 and μ_2 are their respective chemical potentials per concentration unit in that volume element, and μ_1^0 and μ_2^0 are their potentials at an infinite distance. If the solution is infinitely dilute in ions, the average concentration is that at infinite distance from the ion. To solve Eq. (2), we must know c_1, c_2, μ_1, μ_2, D and r as functions of the same variable, which we will take as c_2. All but r are properties of the nonelectrolyte mixture without ions.

At equilibrium there will be no change in the electrostatic chemical potential of the ion if an infinitesimal portion of the second component is moved from any volume element to any other and the corresponding amount of the other component is exchanged for it. That is,

$$(\mu_1-\mu_1^0)dc_1/dc_2+(\mu_2-\mu_2^0)+(\epsilon^2 z_i^2/8\pi r^4)d(1/D)/dc_2=\text{constant} \qquad (3)$$

when $1/r^4=0$, $\mu_1=\mu_1^0$ and $\mu_2=\mu_2^0$, so the constant is zero. Moreover,

$$dc_1/dc_2=-\bar{v}_2/\bar{v}_1 \qquad (4)$$

if $\bar{v}_1$ and $\bar{v}_2$ are the partial volumes per concentration unit of the two components in the volume

* Presented at the Symposium on the Structure of Molecules and Aggregates of Molecules at the Fifth Annual Symposium of the Division of Physical and Inorganic Chemistry of the American Chemical Society, Columbia University, New York, December 30, 1940 to January 1, 1941.

[1] P. Debye and J. McAulay, Physik. Zeits. **26**, 22 (1925).
[2] M. Born, Zeits. f. Physik **1**, 45 (1920).
[3] P. Debye, Zeits. f. physik. Chemie **130**, 56 (1927).

element ds. So

$$(\mu_1-\mu_1^0)\bar{v}_2/\bar{v}_1-(\mu_2-\mu_2^0) = (\epsilon^2 z_i^2/8\pi r^4)d(1/D)/dc_2. \quad (5)$$

Equation (5) gives the required relation between r and c_2. To integrate Eq. (2) it is convenient to express dS as $-4\pi r^4 d(1/r)$ to give

$$E_i=\int_{1/r=0}^{1/b_i}\{4\pi r^4[c_1(\mu_1-\mu_1^0)+c_2(\mu_2-\mu_2^0)] +\epsilon^2 z_i^2/2D\}d(1/r). \quad (6)$$

In solutions which are very dilute in the component of lower dielectric constant, which we will specify as the second, both Raoult's law and Henry's law may be assumed to hold and the reciprocal of the dielectric constant, may be expressed as

$$1/D=1/D_1+\delta c_2 \quad (7)$$

with δ a constant. Then $\mu_2-\mu_2^0=RT\ln x_2/x_2^0$ and $\mu_1-\mu_1^0=RT\ln x_1/x_1^0$, if x_1 and x_2 are the respective mole fractions. The latter is relatively negligible, so

$$x_2/x_2^0=c_2/c_2^0=e^{-B^4/r^4}, \quad (8)$$

$$B^4=\epsilon^2 z_i^2\delta/8\pi RT. \quad (9)$$

In the very dilute solution we may replace

$$\ln x_1/x_1^0=\ln(1-x_2)/(1-x_2^0)$$

by

$$(x_2^0-x_2)=\bar{v}_1(c_2^0-c_2),$$

and replace $c_1=(1-\bar{v}_2c_2)/\bar{v}_1$ by $1/\bar{v}_1$. Then Eq. (6) becomes

$$E_i=\int_{1/r=0}^{1/b_i}\{4\pi r^4 RT[(c_2^0-c_2)+c_2\ln c_2/c_2^0] +(\epsilon^2 z_i^2/2)(1/D_1+\delta c_2)\}d(1/r). \quad (10)$$

Substituting the value of c_2/c_2^0 from 8 in 10 yields

$$E_i=\int_{1/r=0}^{1/b_i}\{4\pi r^4 RTc_2^0[1-e^{-B^4/r^4}-(B^4/r^4)e^{-B^4/r^4}] +[(\epsilon^2 z_i^2/2)(1/D_1+\delta c_2^0 e^{-B^4/r^4})]\}d(1/r). \quad (11)$$

The first term in the second bracket may be integrated directly to give the contribution in the pure first component, and the rest of this bracket cancels the third term in the first bracket by Eq. (9), so

$$E_i=\epsilon^2 z_i^2/2D_1b_i+4\pi RTB^3c_2^0 \times\int_{y=0}^{B/b_i}[(1-e^{-y^4})/y^4]dy$$

$$=\epsilon^2 z_i^2/2D_1b_i+(\epsilon^2 z_i^2\delta c_2^0/2B) \times\int_{y=0}^{B/b_i}[(1-e^{-y^4})/y^4]dy. \quad (12)$$

Equation (12) agrees with that obtained by Debye[2] by considering the salting out as that amount of solute displaced due to the electrostatic action of the ion except that Debye's definition of the change of dielectric constant with the concentration is less general than the present one and the symbols have been altered to correspond to our usual practice.

For small values of B/b_i, the value of the integral is B/b_i and Eq. (12) becomes

$$E_i=\epsilon^2 z_i^2/2D_1b_i+\epsilon^2 z_i^2\delta/2b_i=\epsilon^2 z_i^2/2Db_i \quad (13)$$

which is identical with Eq. (1). For large values B/b_i the integral approaches 1.209. For very small ions, then,

$$E_i=\epsilon^2 z_i^2/2D_1b_i+\epsilon^2 z_i^2\delta\ 1.209/2B \quad (14)$$

and the change from the solvent corresponds to that calculated by Eq. (1) for an ion of size $B/1.209$. The series expansions for small and large values of B/b are, respectively,

$$\int_{y=0}^{B/b}[(1-e^{-y^4})/y^4]dy=\frac{B}{b}\left[1-\frac{1}{2!}\frac{1}{5}\left(\frac{B}{b}\right)^4 +\frac{1}{3!}\frac{1}{9}\left(\frac{B}{b}\right)^8-\frac{1}{4!}\frac{1}{13}\left(\frac{B}{b}\right)^{12}+\cdots\right], \quad (15)$$

$$\int_{y=0}^{B/b}[(1-e^{-y^4})/y^4]dy=1.209-\frac{1}{3}\left(\frac{b}{B}\right)^3 +\left(\frac{b}{B}\right)^3 e^{-(B/b)^4}\left[\frac{1}{4}\left(\frac{b}{B}\right)^4-\frac{1}{4}\frac{7}{4}\left(\frac{b}{B}\right)^8+\cdots\right]. \quad (16)$$

This integral is illustrated by the broken lines of Fig. 2.

Since B is proportional to $(\epsilon z_i)^{\frac{1}{2}}$, the limiting value of the electrostatic contribution, given by Eq. (14), increases only as the three-halves power of the valence. Equation (1) gives the

total change of E_i from one pure component to the other, at least by the assumptions that we have made regarding the mixtures, as proportional to the difference in $1/D$, and Eq. (12) gives the initial slope of E_i *versus* $1/D$. For large ions the initial slope is the same as the average slope, but as $(z_i)^{\frac{1}{2}}/b_i$ increases, the initial slope increases much less rapidly and finally reaches the limiting values given by Eq. (14). It is our present problem to determine the intermediate course of E_i *versus* $1/D$.

The theory of the electrostatic contribution to the effect of the nonelectrolyte medium on the rate of reaction between ions is simpler than for the case of uniform media provided that the sizes are small enough so that E_i is practically the limiting value for both of the reactants and for the critical complex, for in this case any lack of symmetry of the charges or the repulsive forces is ineffective and there is no parameter characteristic of the reaction other than the valences of the reactants. Then

$$\ln k_1/k=[|(z_a+z_b)|^{\frac{1}{2}}-|z_a|^{\frac{1}{2}}-|z_b|^{\frac{1}{2}}] \times\epsilon^2(1/D-1/D_1)/2RTB_I, \quad (17)$$

in which k is the rate of reaction in the mixture of dielectric constant D, k_1 is the rate in the component of higher dielectric constant D_1; B_I is the limiting radius for a univalent ion in that mixture, and z_a and z_b are the valences of the reactants. Since the apparent distance of closest approach of the reactants,[4] is given by the relation

$$\ln k_1/k=\epsilon^2 z_a z_b/RT\bar{R}, \quad (18)$$

the apparent lower limit of $\bar{R}$ is $\bar{R}_\infty$

$$\bar{R}_\infty=B_I 2z_a z_b/[|(z_a+z_b)|^{\frac{1}{2}}-|z_a|^{\frac{1}{2}}-|z_b|^{\frac{1}{2}}]. \quad (19)$$

The values of $\bar{R}_\infty/B_I$ for various types of reaction are given in Table I.

TABLE I. *Apparent limiting distances of closest approach for reaction rates.*

z_a	z_b/z_a	$\bar{R}_\infty/B_I$	z_b/z_a	$\bar{R}_\infty/B_I$
1	1	2.42	−1	1.00
2	1	3.41	−1	1.41
3	1	4.18	−1	1.73
1	2	2.92	−2	1.41
1	3	3.33	−3	1.78
2	3/2	3.80	−3/2	1.71

[4] G. Scatchard, Chem. Rev. **1**, 229 (1932); J. Chem. Phys. **7**, 25 (1939).

If the valences of the two reactants have the same sign, the limit for the complex might be expected to be independent of the distribution of the charges within it. When $z_b=-z_a$, however, any electrostatic effect depends entirely upon this distribution. Equation (19) and Table I represent superposition of the charges. A thorough study of this case would require calculation of the sorting of solvent molecules around a dipole, which is a problem much more difficult than the one attacked here. We can be sure that the sizes in the last column of Table I are lower limits which are much too small. If the behavior of a dipole molecule made up of two ions of equal size has the same relation to that of one of the ions as in a homogeneous solvent, this limit should be twice as large. Qualitative reasoning shows that it might be even larger than this, and the most probable value of the limiting size for any reaction between ions of opposite sign may well be the same as for the reaction between ions with the same absolute values of the valence, but with the two valences of the same sign. If so, the values in the third column of Table I should be used for the valence ratio in the fourth column as well as for that in the second.

For water-ethanol mixtures we have taken from the measurements of Wyman[5]

$$D=78.54-40.32n_2-25.68n_2^2+11.74n_2^3, \quad (20)$$

in which n_2 is the volume of alcohol in unit volume of the mixture. From the measurements of Dobson[6] we have taken as before[7]

$$\mu_1=2.303RT[\log x_1+0.9(1-x_1)^2-0.5(1-x_1)^3],$$
$$\mu_2=2.303RT[\log (1-x_1)+0.15x_1^2+0.5x_1^3]. \quad (21)$$

The partial volumes were obtained by numerical differentiation of the volume measurements of the National Bureau of Standards[8] without troubling to obtain an analytical expression. For the limiting case of infinitesimal ethanol concentration, B is 1.89A, the limiting apparent diameter is 1.56A, and B/b_i should exceed unity for many real ions. Because of the fourth power in Eq. (8), the ratio c_2/c_2^0 decreases very rapidly when B/b_i becomes greater than unity. The

[5] J. Wyman, Jr., J. Am. Chem. Soc. **53**, 3292 (1931).
[6] H. J. E. Dobson, J. Chem. Soc. **127**, 2866 (1925).
[7] G. Scatchard, J. Am. Chem. Soc. **49**, 217 (1927).
[8] N. S. Osborne, E. C. McKelvy and H. W. Bearce, Bull. Bur. Stand. **9**, 327 (1913).

ratio is eight ten-thousandths when $r/\sqrt{z_i}$ is 1A.

For all compositions the important characteristic of the ion is $z^{\frac{1}{2}}/r$. Although the exact relationships are very complicated and depend upon the properties of all solutions less concentrated than the average being considered, our computations serve to show which generalizations of the limiting case may be extended as approximations to the higher alcohol concentrations. For the limiting case B^4 may also be written

$$B^4=\epsilon^2 z_i^2 v_1(1/D_1-1/D)/8\pi\mu_1 \qquad (22)$$

since $RT=-\mu_1/x_2=-\mu_1/c_2v_1$ if the pure first component is chosen as the standard state in which $\mu_1=0$, and if $\delta=(1/D-1/D_1)/c_2$. With this definition Eq. (12) gives a fair approximation of the chemical potential for all compositions. We may obtain an approximate value for the concentration if we replace the concentration ratio of Eq. (8) by the dimensionless distribution coefficient to give

$$c_2c_1^0/c_1c_2^0=x_2x_1^0/x_1x_2^0=e^{-B^4/r^4} \qquad (23)$$

and use the value of B from Eq. (22).

Computations were made for mixtures containing 25, 50, 75, 90, 95 and 99 weight percent alcohol, for which the approximate mole percentages and approximate percentage changes of reciprocal dielectric constant, $1/D-1/D_1$, are 10, 27, 55, 80, 90 and 98.[9] Figure 1 shows the compositions, expressed as volume fractions of alcohol (the volume of alcohol divided by the volume of alcohol plus that of water) as a function of $r/\sqrt{z_i}$. The full lines are calculated by Eq. (5) and the dotted lines by Eq. (23). These dotted lines show too small changes for large values of $r/\sqrt{z_i}$, and too large changes in the intermediate range of rapid change. For small values of $r/\sqrt{z_i}$ the two equations give the same rate, but a logarithmic scale is necessary to show the change in this region. Probably the most important relation to be noted in this figure is that the concentration change is so very large when $r/\sqrt{z_i}$ is about unity, so that for $r/\sqrt{z_i}$ less than 1A the medium is practically pure water even for the most concentrated alcohol solution, and that most of the concentration change occurs between values of the abscissa of 1 and 3A. When $r/\sqrt{z_i}$ is 5A the concentration is practically the same as when it is infinite.

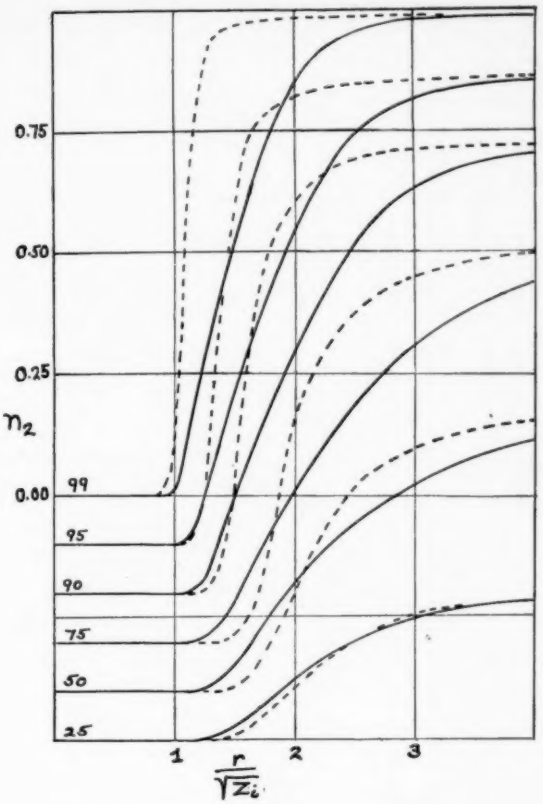


FIG. 1. Concentration of ethanol around ion.

Figure 2 shows the effect on the free energy of the ion, expressed as

$$[E_i-\epsilon^2 z_i^2/2D_1b_i]/[\epsilon^2 z_i^2(1/D-1/D_1)/2\sqrt{z_i}]$$

versus $(z_i)^{\frac{1}{2}}/b_i$. The full line of unit slope represents Eq. (1) and the asymptote for large sizes; the full curve and horizontal line represent Eq. (6) and its asymptote for small sizes. The broken curve and horizontal line represent the corresponding values for the approximate Eqs. (12) and (22). Where they are omitted they practically coincide with the corresponding full lines. The shapes of all the curves are closely the same, and any one could be closely approximated by Eq. (12) if B were determined from the horizontal asymptote. Since this asymptote can not be determined without knowledge of the rest of the curve, however, this has little advantage except to simplify the present discussion.

[9] Similar calculations have been made by J. McAulay, J. Phys. Chem. **30**, 1202 (1926), assuming that water-ethanol mixtures are ideal solutions.

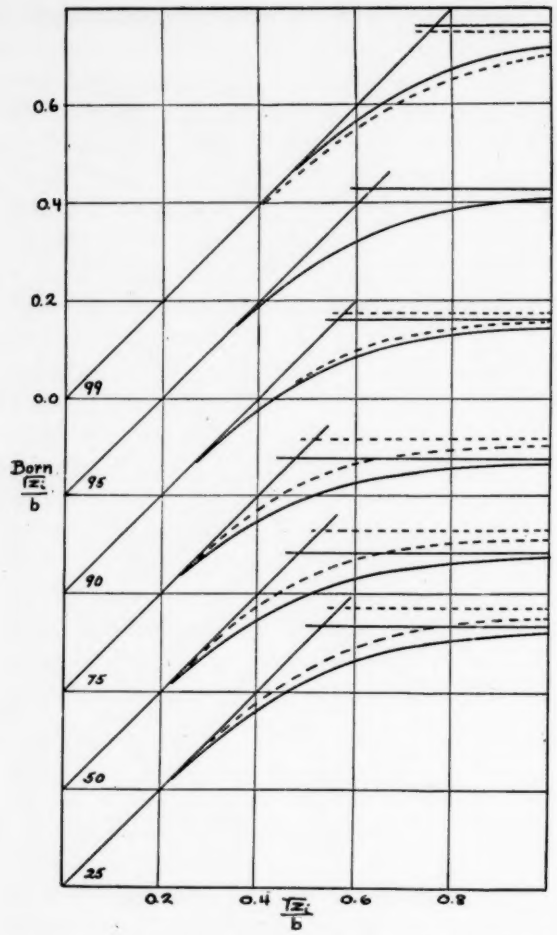


FIG. 2. Electrical free energy as function of ion size.

When $(z_i)^{\frac{1}{2}}/b_i$ is less than half the horizontal asymptote, the curve is practically on the first asymptote; when the abscissa is greater than twice the horizontal asymptote, the curve is practically on this second asymptote. Where the two cross, the curve is 84.5 percent of its asymptotes, and it is never less than 84.5 percent of the smaller asymptote.

The value of the horizontal asymptote, $1.209/B$ decreases with increasing alcohol concentration to a minimum between 25 and 50 weight percent alcohol, which is about three-fourths of the initial value. It then increases toward infinity for pure alcohol, but only reaches five-thirds of the minimum value in 99 percent alcohol. The approximate equation has a slightly larger minimum and the subsequent increase is somewhat slower, so that the two curves cross at about 95 percent alcohol. Table II contains the values of $1.209/B$ given by the two equations, as well as the compositions expressed in various ways and the relative change of the reciprocal dielectric constant. It will be noted that this latter quantity never varies more than 2 percent from the mole fraction of the alcohol.

Figure 3 shows

$$[E_i-\epsilon^2 z_i^2/2D_1b_i]\Big/\frac{\epsilon^2 z_i^2}{2}\left(\frac{1}{D_2}-\frac{1}{D_1}\right)$$

plotted against $(1/D-1/D_1)/(1/D_2-1/D_1)$ for ethanol-water mixtures. The two straight lines correspond to Eq. (1) for $b_i/\sqrt{z_i}=1.33$A. The full curves correspond to Eq. (6), and the broken curves to Eqs. (12) and (22). The upper set gives the limit for small radii, and the lower set gives the values for $b_i/\sqrt{z_i}=1.33$A. The differences between the approximate equations and the complete calculation are small relative to the differences of either from Eq. (1).

For the total chemical potential of the electrolyte we write

$$\mu_s-\mu_{s0} = \Sigma_i \nu_i[RT\ln x_i/x_{i0}+E_i-\epsilon^2 z_i^2/2D_1b_i], \quad (24)$$

in which ν_i is the number of ions of the ith type corresponding to a molecule of the electrolyte, and the subscript zero refers to the standard state. For small ions the squeezing out of the non-electrolyte of lower dielectric constant from around the ion makes the short range forces acting on the ion almost independent of the non-electrolyte concentration and thus increases the accuracy of the assumption that the solutions are ideal. The argument of Bell and Gatty[10] that

TABLE II. *Concentrations and limiting reciprocal radius for ethanol-water at 25°.*

					1.209/B	
w_2	z_2	n_2	x_2	$\frac{1/D-1/D_1}{1/D_2-1/D_1}$	Eq. (6)	Eq. (12)
0.00	0.000	0.000	0.000	0.000	0.639	0.639
.25	.297	.305	.115	.101	.538	.573
.50	.560	.579	.281	.272	.487	.531
.75	.792	.813	.540	.555	.478	.517
.90	.920	.933	.779	.802	.562	.578
.95	.960	.968	.881	.899	.633	.636
.99	.992	.994	.975	.980	.764	.751

w_2 = weight ethanol per unit weight mixture;
z_2 = volume ethanol per unit volume ethanol plus water;
n_2 = volume ethanol per unit volume solution;
x_2 = moles ethanol per mole solution.

[10] R. P. Bell and O. Gatty, Phil. Mag. [7] **19**, 66 (1935).

the norm in such studies should be moles per unit volume rather than mole fraction seems quite untenable. Their model of a "non-interacting solution" which leads to volume concentrations, assumes that the solute has no interaction with the solvent, even of repulsion, other than those specified. The model which leads to mole fractions assumes, perhaps only approximately, that the unspecified interactions of a solute molecule are the same as those of a solvent molecule except for the size of the volume of exclusion due to the forces of repulsion. Although it does not make a satisfactory starting point for a thorough-going physical treatment in which the short range and repulsive fields are included, it is certainly much to be preferred for a treatment which specifies only electrostatic fields.

Comparison with experimental results shows that this treatment raises at least as many difficulties as it settles. The straight lines of Fig. 3 for Eq. (1) with $b_i/\sqrt{z_i}=1.33$A correspond closely to the values calculated from the measurements of the solubilities of sodium and potassium chlorides.[11] The results are always higher than those calculated by Eq. (6) for this value of b_i, and are higher than the limiting value except at very high alcohol concentrations. It is extremely unlikely that the discrepancy can be attributed to errors in the experimental values. The measurements were made by several independent observers, which should leave small chance for systematic errors. Although they are corrected for the ratio of the activity coefficients in saturated solutions to those at infinite dilution in the same medium by the Debye-Hückel approximation, any effect of the higher terms would be to increase the values at higher alcohol concentrations, and so to make the discrepancy larger.

We might expect another effect on the extrapolation to infinite dilution because of our assumption that the average concentration is the same as the concentration at infinite distance from the ion. In a solution whose total ion concentration is one mole per liter, the volume per ion is 1000/0.6 cubic angstroms. If we take the maximum value of r as half the cube edge which corresponds to this volume, it is about 6A. Integration of the total quantities of the two components within a fixed radius show that the difference between the average concentration and that at infinite distance is small for 6A and quite negligible for twice this distance, which corresponds to one-eighth molal total ion concentration. So this should not give a measurable error.

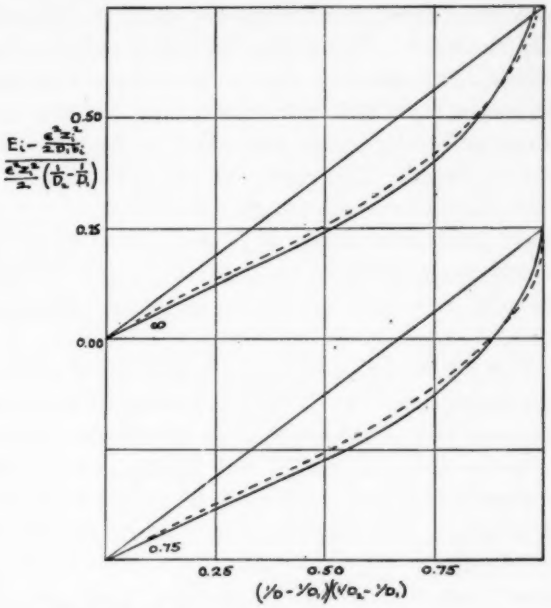


FIG. 3. Electrical free energy as function of reciprocal dielectric constant.

The difficulty of this treatment is further brought out by a consideration of dilute solutions of solutes which increase the dielectric constant considerably. There will usually not be sufficient information to apply Eqs. (5) and (6), especially if the solute is a solid, but Eqs. (12) and (19) may be applied with the subscript one attributed to the solute.[12] Although the maximum change in the reciprocal dielectric constant is only half of that for ethanol-water mixtures, the molecular volume of the solvent will usually be more than twice that of water. The value of B^4 will change rapidly with the concentration and will not be very different from that of ethanol-water mixtures in which the

[11] G. Scatchard, J. Am. Chem. Soc. **47**, 2098 (1925); Chem. Rev. **4**, 383 (1927); Trans. Faraday Soc. **23**, 454 (1927).

[12] It is, of course, quite incorrect to apply these equations with water as the solvent and using a negative value of δ as is done by W. J. Dunning and W. J. Shutt, Trans. Faraday Soc. **34**, 1192 (1938).

mole fraction of the component of higher dielectric constant is the same. So the apparent size should decrease as the solute concentration decreases, and for dilute solutions should be considerably less than the real size. The experimental results, however, indicate an apparent size independent of the concentration and not appreciably smaller than those for substances which decrease the dielectric constant, which should by our treatment show too large apparent sizes.

Dioxane-water mixtures show a quite different sort of difficulty. The freezing point measurements of Scatchard and Benedict[13] show excellent agreement with Eqs. (12) and (22) for potassium, sodium and lithium chlorides in dilute solutions of dioxane. The measurements of Davis, Ricci and Sauter[14] of the solubilities of silver acetate, silver sulfate and barium iodate also appear to agree in solutions dilute in dioxane, but in more concentrated solutions the curve of log S against $1/D$ is nearly flat. Table III shows the values of $1/b$ determined by Eq. (1) from their measurements. These are the reciprocals of "a calc. based on mole fractions" of their Tables XI, XII and XIII except that the a values for barium iodate are multiplied by

$$(\log x p_{H_2O}/p^0_{H_2O})/\log x(p_{H_2O}/p^0_{H_2O})^{\frac{1}{2}}$$

to take into account the fact that the water activity should multiply the solubility product rather than the solubility itself. The second and fifth columns give the maximum values of $1/b_i$ calculated by Eqs. (12) and (22) together with the dielectric constant measurements of Åkerlöf and Short[15] and the vapor pressure measurements of Hovorka, Schaeffer and Dreisbach.[16] The second column is for a uni-univalent electrolyte such as silver acetate, and the fifth column is for a uni-bivalent or a bi-univalent electrolyte such as silver sulfate or barium iodate. The explanation of these results is complicated by the further fact that the activity coefficient increases so rapidly between 95 and 100 percent dioxane that the solubilities of these salts are immeasurably small in pure dioxane.

TABLE III. *Reciprocal radii in dioxane-water mixtures at 25°.*

Wt. Fraction Dioxane	Max. 1–1	Silver Acetate	Silver Sulfate	Barium Iodate	Max. 1–2, 2–1
0.1	0.387	0.242	0.254	0.275	0.311
.2	.382	.229	.236	.251	.307
.3	.370	.218	.218	.230	.298
.4	.365	.204	.197	.201	.294
.5	.355	.182	.175	.161	.286
.6	.340	.157	.147	.121	.274
.7	.325	.141	.107	.073	.262
.8	.307	.089	.068	.044	.247
.9	.290	.058	.041	.035	.233

The present treatment neglects the effect of electrical saturation, as noted by Debye,[3] and the discrete structure of the nonelectrolyte mixture, as noted by Scatchard and Benedict.[13] The first effect depends upon the relative saturation of the nonelectrolyte components, and even its sign is difficult to predict. It is probably much smaller than the second effect. The volume occupied by an ethanol or dioxane molecule corresponds to an average diameter of 5–6A. Even if one edge of the molecule reaches to the edge of an ion of 1A radius, more than half of the molecule is at a distance greater than 3A. Our treatment says that there should be almost no solute at 1A, regardless of the average composition, but that the composition at 3A should be almost the same as the average composition. The concentration of tangent molecules must depend upon an integral over the whole molecule and correspond to what our theory would give for a considerably greater distance, whose exact magnitude would depend upon the shape of the molecule and upon the dielectric constant of its various parts.

Calculations for very dilute solutions of a solute with spherical molecules of radius b_n and spherically symmetrical dielectric constant show that, when b_i is large relative to b_n, the effect is the same as for Eq. (1) or (6) for a size (b_i+b_n), and that as b_i decreases, the effect increases more rapidly than for Eq. (1) for (b_i+b_n) but less rapidly than even Eq. (6) for b_i. It is easy to see that a nonelectrolyte molecule with an unsymmetrical dielectric constant will give an effect qualitatively similar to that of a smaller symmetrical molecule, but quantitative calcula-

[13] G. Scatchard and Marjorie A. Benedict, J. Am. Chem. Soc. **58**, 837 (1936).

[14] T. W. Davis and J. E. Ricci with C. G. Sauter, J. Am. Chem. Soc. **61**, 3274 (1939).

[15] G. Åkerlöf and O. A. Short, J. Am. Chem. Soc. **58**, 1241 (1936).

[16] F. Hovorka, R. A. Schaeffer and D. Dreisbach, J. Am. Chem. Soc. **59**, 2753 (1937).

tion of the effect of lack of symmetry would be complicated. A detailed consideration of nonelectrolytes of finite size will be given in a later paper. One danger which this treatment accentuates should be noted here—the danger of considering the nonelectrostatic part of the chemical potential as a function only of the concentration, independent of the concentration gradient even though the change is enormous over the dimensions of a single molecule, and independent of any order in the orientation even though the orientation must be nearly fixed.

We may conclude that the extension of the Debye method which takes into account the sorting of nonelectrolytes near an ion but neglects their discrete structure represents the solubilities of sodium and potassium chlorides in ethanol-water mixtures less well than the simple Born expression, and fails to represent the flattening of the solubility curves in dioxane-water mixtures or the similar flattening of the reaction rate curves in many mixtures. If the size of the nonelectrolyte molecules is to explain the flattening, the effect of size must be much larger for solutions rich in dioxane than for those containing a small amount; if it is also to explain the solubilities in ethanol, asymmetry must play an important role.

JANUARY, 1941 JOURNAL OF CHEMICAL PHYSICS VOLUME 9

The Effect of the Rotation of Groups about Bonds on Optical Rotatory Power*

WALTER KAUZMANN AND HENRY EYRING
Frick Chemical Laboratory, Princeton University, Princeton, New Jersey

(Received October 9, 1940)

It is shown that the numerical value of the optical activity of an optically active compound is markedly reduced if the groups surrounding the asymmetric atoms in the compound possess a threefold axial symmetry about the bonds connecting them with the asymmetric atoms. This threefold symmetry may be either inherent in the groups themselves, or it may be acquired by them through the free rotation of the groups about these bonds or by their orienting themselves to equal extents in each of the three possible equilibrium positions about each bond. The order of magnitude of the numerical value of the optical activity is thus a measure of the freedom of orientation about single bonds of the groups in an asymmetric molecule. That this effect is actually a dominant factor in determining the order of magnitude of the optical activity is proved by the contrast in the optical activities of cyclic and open chain compounds. It is shown to have an important effect on the temperature coefficient of optical activity and to lie at the root of the observed differences in the orders of magnitude of the rotatory powers of liquids and crystalline solids. The structures of certain sulphur compounds, polypeptides, and proteins are discussed in the light of the magnitudes of their optical activities.

A STUDY of the dependence of the optical rotations of many substances upon the wave-length of the light used in their measurement has shown that rotatory power is closely associated with the absorption bands characteristic of the groups in the molecule. This has given rise to the concept of a *chromophoric group* whose electronic transitions have been rendered optically active by the *vicinal actions* of surrounding groups.[1] It is the purpose of this paper to show that, without going into the detailed physical nature of these vicinal actions, a consideration of only certain of their very general and fundamental characteristics makes it possible to use the optical rotatory power as a tool in investigating molecular structure and conformation.

When one proceeds to a calculation of the optical activity of a molecule, the vicinal actions which are to be considered can be classified

* Presented at the Symposium on the Structure of Molecules and Aggregates of Molecules at the Fifth Annual Symposium of the Division of Physical and Inorganic Chemistry of the American Chemical Society, Columbia University, New York, December 30, 1940 to January 1, 1941.

[1] Kuhn, Trans. Faraday Soc. **26**, 293 (1930).

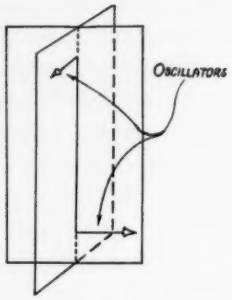


FIG. 1.

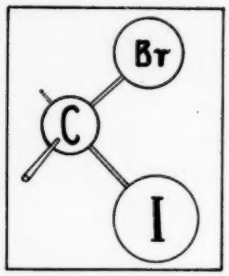


FIG. 2.

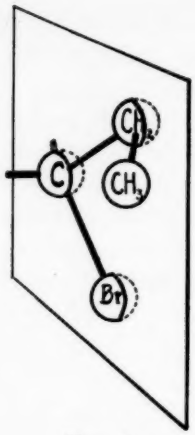


FIG. 3.

naturally according to a general order as follows. A chromophoric group such as OH or Br, when it is separated from the influences of neighboring groups, is not optically active since it possesses a plane of symmetry. Under certain circumstances, a *single* neighboring group may, however, distort the chromophore in such a way as to destroy its symmetry and render it optically active. Such a vicinal action we call a *first-order vicinal action*. The sum of all of the separate vicinal actions of individual groups on otherwise symmetrical chromophores we may speak of as giving rise to the *first-order contribution* to the optical rotation. The next type of vicinal action arises from the interaction of one vicinal group with a chromophore which has already been considered as distorted by another group. This gives rise to a *second-order vicinal effect* and a *second-order contribution* to the optical rotation. Similarly, we may have vicinal effects and contributions of third and higher order.

To illustrate a first-order vicinal action we may take Kuhn's simple case of two coupled oscillators:[1] A single linear oscillator is optically inactive, but if another linear oscillator with a suitable interaction and orientation in space be introduced, the system may become optically active (see Fig. 1). A second-order vicinal effect might arise as follows: Two three-dimensional isotropic oscillators when brought close to one another are no longer isotropic. The slight anisotropy thus produced may now interact with a suitably oriented linear oscillator to give rise to optical activity in a manner similar to the case of two linear oscillators.

The simplest way to determine the order of a given vicinal action is to ascertain how many groups are interacting simultaneously in the perturbation of the chromophoric group. Thus, any interaction involving two groups at a time (one vicinal group and one chromophore) will be a first-order vicinal action, while one involving the interaction of three groups (two vicinal groups and one chromophore) will be a second-order vicinal action. The sum total of all groups taken two at a time will give the first-order contribution to the rotation, while the sum total of all interactions of three groups taken at a time will give the second-order contribution to the rotation.

First-order contributions to the rotation should in general be larger than second- and higher-order contributions. This is evident from the examples given above, where the ratio of the first-order contribution to the second-order contribution would be of the order of magnitude of the ratio of the anisotropy of a group such as C : O, in which the atoms are separated by one or one and a half angstroms, to the mutually induced anisotropies of two atoms about two or three angstroms apart. Since the induced anisotropy of two otherwise isotropic spheres varies inversely as the cube of the distance between them, this would signify that in general the first-order contribution should be at least five or ten times as large as the second-order contribution. This estimated ratio agrees with that found by other and more general lines of reasoning.[2] The third- and higher-order contributions

[2] Kauzmann, Walter and Eyring, Chem. Rev. **26**, 339 (1940).

to the rotation will, of course, be by the same reasoning of even less importance than the second and first.

We now consider the results of vicinal actions according to the above classification for some simple types of substances. First, let us consider the interactions of two axially symmetrical groups such as those shown in Fig. 2. It is at once evident that two such groups can alone give rise to no first-order contribution to the optical activity, since there is a plane of symmetry for the three atoms there shown. Next, consider the interaction of an ethyl group and a C−Br bond (Fig. 3). As long as no plane can be drawn through the ethyl group and the C−Br bond, there can be a first-order contribution to the optical activity. But if, in a given compound, the ethyl group is freely rotating about the bond connecting it with the asymmetric carbon atom so that it becomes axially symmetric around this bond, or even if it can take up with equal probability several positions which are, however, symmetrically disposed about the plane shown in Fig. 3, then no first-order contribution to the optical activity is possible for these groups.

These arguments may easily be extended to prove the general statement that *if the groups about an asymmetric atom possess cylindrical symmetry about the bond connecting them to the asymmetric atom, or if the groups can effectively acquire this symmetry by freely rotating or by taking up a number of equilibrium positions, then the first-order contribution to the optical activity will vanish and any observed rotation must be due to higher order effects.* Coupling this statement with what we know about the relative orders of magnitude and of first- and second-order effects, we may state further that *those influences which tend to restrict freedom of rotation and of orientation about bonds will tend to increase the order of magnitude of the optical activity.*

The consequences of this statement will now be considered.

1. Comparison of the rotatory powers of substances in which the asymmetric centers are in closed rings and open chains

Since ring formation severely restricts the freedom of rotation of groups about bonds, it should have a big effect in increasing the optical activity when an asymmetric center is one of the ring atoms. Such an effect has long been noticed[3] and is particularly striking in those substances (e.g., polyhydroxy alcohols and sugars) for which there is the most reason to believe that the open chain form can exhibit the greatest freedom of rotation because the groups attached to the asymmetric carbon are the smallest, as will be seen in Table I, where the rotations of various types of open chain and closed ring compounds are compared.

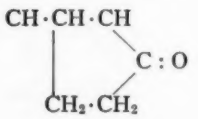
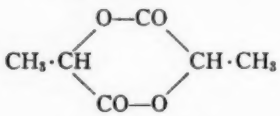
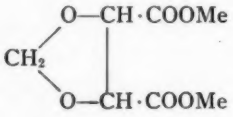
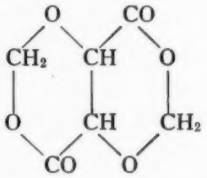
This very low optical activity of many open chain compounds has a bearing on the question of the conformations assumed by long chains in general. It has been suggested by some workers that such chains are usually stretched out, while others believe that chains tend to roll up into balls; from the evidence presented here, however, it would seem that in the absence of large side groups, chains are predominantly neither stretched out nor rolled up into balls, but that they tend to take up all conformations with very nearly the same probability. Recent theories of the elasticity of rubber have assumed this randomness to occur even in the absence of a solvent, and some evidence as to the validity of this assumption could be found in the rotations of glasses of polyhydroxy alcohols. A persistence of the low rotations of these substances as glasses would indicate that their chains are randomly twisted even when no small solvent molecules are present to fill up the holes in the structure of the glass which result from the poor packing of the molecules as a consequence of this randomness.

2. Effect of temperature on optical activity

Changing the temperature will affect the optical activity of a substance chiefly in two ways: by changing the extent of the interaction the active molecules and the solvent, and by changing the conformations of the active molecules. It will be shown later that an increase in the temperature generally causes the interaction between solvent and solute to decrease, so that at higher temperatures the optical activity of a substance tends to become independent of the solvent in which it is dissolved

[3] Van't Hoff, *Arrangement of Atoms in Space* (Longmans, New York, 1898), p. 146.

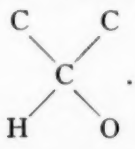
TABLE I. *Comparison of the rotations of cyclic and open chain compounds.*

OPEN CHAIN COMPOUNDS		CYCLIC COMPOUNDS	
Type $CH_2OH \cdot (CHOH)_n \cdot CH_2OH$	$[M]_D$	Type $H \cdot (CHOH)_n \cdot CH \cdot (CHOH)_3 \cdot CHOH$ (O-bridge)	$[M]_D$
Arabite	8	Ribose	32
Talite	6	Alpha-arabinose	263
Mannite	1	Beta- "	75
Idite	6	Alpha-xylose	138
Sorbite	4	Alpha-mannose	54
		Beta- "	25
		Alpha-galactose	260
		Beta- "	25
		Alpha-glucose	302
		Beta- "	34
Type $CH_2OH \cdot (CHOH)_4 \cdot CH(OMe)_2$		Type $CH_2OH \cdot CH \cdot (CHOH)_3 \cdot CHOMe$ (O-bridge)	
Mannose dimethyl acetal	1	Alpha-methyl mannoside	159
Galactose " "	34	Alpha-methyl galactoside	345
Glucose " "	34	Beta- " "	1
		Alpha- " glucoside	307
		Beta- " "	62
Type $CH_2OAc \cdot (CHOAc)_4 \cdot CH(OMe)_2$		Type $CH_2OAc \cdot CH \cdot (CHOAc)_3 \cdot CHOMe$ (O-bridge)	
Mannose dimethyl acetal pentacetate	69	Alpha-methyl mannose tetracetate	162
Glucose " " "	52	Beta- " " "	170
		Alpha- " galactose "	474
		Beta- " " "	65
		Alpha- " glucose "	478
		Beta- " " "	50
$CH_3 \cdot CH_2 \cdot CHCH_3 \cdot CH_2 \cdot CHO$	9		130
$CH_3 \cdot CHOH \cdot COOH$	$2[M]_D = 9$		$[M]_D = 400$
$CH_3CHOH \cdot COOEt$	$2[M]_D = 29$		$[M]^{100°C}_{5461} = 204$
$COOEt \cdot CHOH \cdot CHOH \cdot COOEt$	$[M]^{190°C}_{5461} = 35$		$[M]^{20°C}_{5461} = 129$

ACID	$[\alpha]_D$ OF FREE ACID	$[\alpha]_D$ OF GAMMA-LACTONE	$[\alpha]_D$ OF DELTA-LACTONE
Arabonic	18	74	—
2,3,4 trimethyl arabonic	16	—	178
2,3,5 " "	2	44	—
2,3,4 " lyxonic	13	—	35
2,3,5 " "	21	82	—
3,4 dimethyl rhamnonic	16	—	153
Rhamnonic	8	40	98
Gluconic	7	68	62
Mannonic	10	52	112
Galactonic	13	77	—
2,3,4,6 tetramethyl gluconic	22	—	72
2,3,5,6 " "	34	101	—

(aside from a factor of $(n^2+2)/3$, where n is the refractive index). An increase in the temperature generally affects the conformation of the molecule by increasing the freedom of rotation about bonds—and if there is not complete freedom of rotation, due to the presence of a potential barrier of the type supposed to exist in ethane, then an increase in the temperature will tend to equalize the times spent by a group in the various positions of minimum potential energy on either side of the barriers. In either case, when an symmetric center is not a member of a closed ring, a temperature increase brings about more and more axial symmetry of the groups attached to the asymmetric atom, and so reduces the first-order contribution to the optical activity. From what we have said before, it is evident that this means that at high temperatures all open chain compounds should have rather small rotations, while the orders of magnitude of the optical activities of compounds in which the asymmetric centers are the members of rings should be unaffected.

Before proceeding to an investigation of the actual data, however, it will be convenient to study some further structural effects which will influence the magnitudes of higher order contributions to the rotatory power. Consider the secondary butyl alcohol molecule ($CH_3 \cdot CH_2 \cdot \cdot CHMe \cdot OH$). Suppose that all the groups attached to the asymmetric atom are axially symmetric due to free rotation or its equivalent, so that the first-order vicinal actions are all zero. The largest of the second-order vicinal actions will clearly be those involving groups and portions of groups closest to the asymmetric carbon atom. Therefore, in secondary butyl alcohol, the most important of the higher order vicinal actions should be those involving the atoms



But this group of atoms possesses a plane of symmetry, so cannot by itself contribute to the optical activity. Nonvanishing higher order vicinal actions in secondary butyl alcohol at high temperatures only become important when the terminal methyl of the ethyl group is considered along with the rest of the molecule. Now this methyl group is some distance away from the asymmetric center, so that its interactions with the other groups will be rather weak, and the higher order contributions to the optical activity in secondary butyl alcohol should be much smaller than those in, say,

$$CH_3{-}\underset{H}{\overset{Br}{C}}{-}I, \text{ or } CH_3{-}\underset{H}{\overset{Br}{C}}{-}CN, \quad CH_3{-}\underset{H}{\overset{OH}{C}}{-}COOH,$$

where all of the atoms directly attached to the asymmetric center are considerably different from one another. (It is to be expected that a carbon atom in a COOH group or a $-CN$ will behave optically in a very different way from one in a CH_3 group, since it has been very differently modified by its neighbors.)

In what follows, therefore, we shall divide the compounds which we consider into various classes. In one class (referred to hereafter as *class A*) we shall have those substances in which axial symmetry of groups attached to asymmetric centers is rendered impossible due to the presence of a ring in the molecule. Another class of substances (*class B*) will have groups which rotate freely about the bonds attaching them to the asymmetric centers and will have two or more similar atoms attached to each of these centers. Secondary butyl alcohol would thus fall in class *B*. A third class (*class C*) will include those substances in which axial symmetry of groups is possible, but in which the atoms directly attached to the asymmetric center are considerably different in their optical behavior. Lactic acid therefore belongs in class *C*.

Class *A* compounds should show large optical rotations at both high and low temperatures, and if a class *A* compound has a large optical activity, an increase in the temperature should be just as apt to cause an increase in the optical activity as it is to cause a decrease. Class *B* and *C* compounds should generally tend to show low optical rotations, but if any should happen to have large rotations, then an increase in the temperature should cause a decrease in the magnitudes of these rotations. Class *C* compounds should generally have larger rotations at high temperatures than do class *B* compounds.

In Figs. 4 and 5 are represented the temperature coefficients of the rotations of most of the

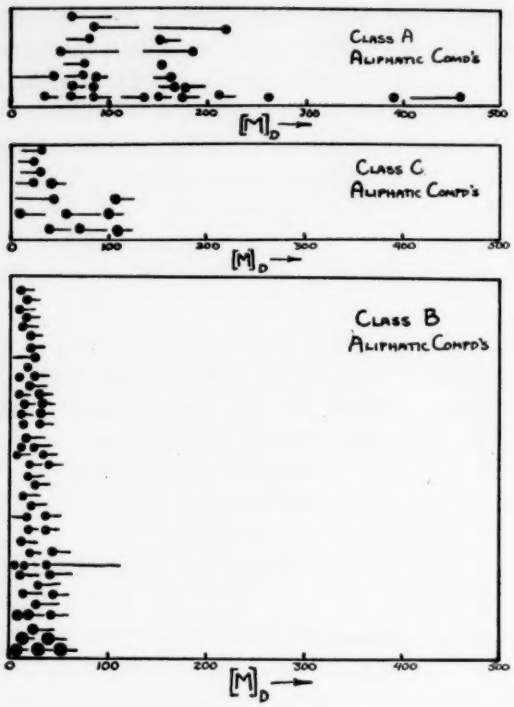


FIG. 4.

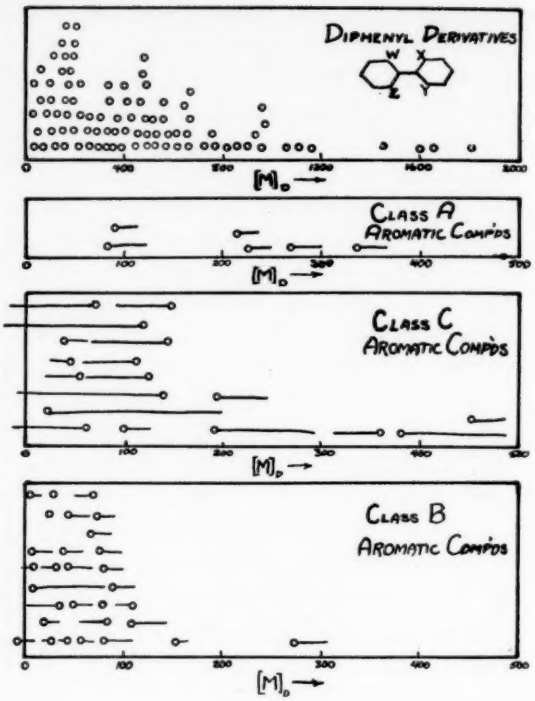


FIG. 5.

substances in the literature for which data are available over an appreciable temperature range. In the figures, a black dot represents the rotation of one substance at the highest temperature at which measurements have been made. From each dot is drawn a line in the direction toward which the rotation tends as the temperature is lowered. The length of each line has been made numerically equal to one hundred times the rate of change of the rotation of each substance per degree centigrade at 20°C. Figure 4 includes only aliphatic compounds, and Fig. 5 includes only aromatic compounds; this distinction has been made in order to keep more constant all of those factors not immediately concerned in the determination of the amount of rotation of groups about bonds.[4]

A study of the figures will show the following:

(*a*) Among the aliphatic compounds the open chain molecules (classes *B* and *C*) with large rotations all tend to have smaller rotations as the temperature increases, while the ring molecules (class *A*) have temperature coefficients whose sign is independent of whether or not their rotations are large or small. This is as we have predicted.

(*b*) Aliphatic compounds of class *B* tend to have smaller rotations at high temperatures than do those of class *C*. This is evident from the fact that the rotations of a larger proportion of the latter tend to increase with increasing temperatures, even when fairly large rotations are involved.

(*c*) Aromatic compounds as a whole possess larger rotations than do aliphatic compounds, so that the relationship between class *B* and class *C* compounds, although similar to that

[4] Most of the data used here are from (1) a long series of papers by Pickard, Kenyon *et al.* in J. Chem. Soc. London, 1911 ff.; (2) Guye and Amaral, Comptes rendus **120**, 1345 (1895); (3) Guye and Aston, Comptes rendus **124**, 194 (1897); and (4) some unpublished data from this laboratory. Where there are data in the literature for a large number of compounds belonging to the same homologous series, only two or three examples have been included in the figures. Thus, Pickard and Kenyon (J. Chem. Soc. **99**, **45** (1911); **101**, 620 (1912); **105**, 830 (1914)) have determined the effects of temperature on eleven alkyl carbinols and 73 of their fatty acid esters. The results among the carbinols are very similar; in all, there is a slight initial decrease in rotation to a small, constant value. At high temperatures there is a very slight tendency for the rotation to increase, probably associated with changing solvent effects as the boiling point is approached. For all of the 73 esters, too, the temperature *vs.* rotation curves are very similar, all being, in fact, very nearly superimposable.

found among aliphatic substances, occurs against a background of an entirely different scale of orders of magnitude. The relationship between aromatic substances of class *A* and those of classes *B* and *C* is obscured by a scarcity of data for the class *A* compounds. To circumvent this partially we have added to Fig. 5 a "map" of the distributions of the orders of magnitude of the rotations of some optically active diphenyl derivatives. These substances should behave in the same way as class *A* compounds, since no planes of symmetry can be drawn through pairs of groups attached to the "asymmetric centers" in the molecules. That they do indeed behave like class *A* compounds is shown by their tendency to have very large rotations; so great is this tendency that the scale of rotations for these compounds in Fig. 5 is reduced by fourfold as compared with the scales for the other aromatic compounds in order to make it possible to include the data for all of the known compounds. The dependence on the temperature for these compounds has not been studied, but we would predict that in the absence of racemization, just as many will be found to have rotations whose numerical value increases with the temperature as those whose numerical value decreases with the temperature. Furthermore, this will tend to be just as true among those substances having large rotations as among those having small rotations.

The fact that aromatic compounds seem to show larger rotations than aliphatic compounds may be due to either of two effects. There may be incomplete axial symmetry in even the high temperature forms of the aromatic compounds due to greater steric repulsions by aryl groups, or the effect may be due to a greater polarizability or susceptibility to distortion possessed by aryl groups which tends to increase both the first- and second-order vicinal actions in molecules in which they occur.

(*d*) There would appear to be some relationship between the sizes of groups attached to the asymmetric center and the magnitude of the temperature effect. This shows itself best in the contrast between the magnitudes of the temperature coefficients (lengths of the lines in Fig. 5) for aromatic compounds of classes *B* and *C*, where the direct attachment of the bulky phenyl or naphthyl group to the asymmetric center greatly increases the temperature coefficient of the optical rotation. All of this effect cannot be directly due to the sizes of the groups concerned, however, for in benzoin and in benzoin methyl ether,[5] where there are very large temperature coefficients, not only the magnitude of the temperature effect, but its sign as well depends upon the nature of the solvent. Therefore, the solvent must play a considerable, if not a dominant role in determining many of the temperature coefficients which have been measured. (Of course, when a complex aromatic compound is in the pure, so-called "homogeneous" state, solvent effects, i.e. effects due to surrounding molecules, are every bit as large as when it is dissolved in, say, benzene.) Therefore, only when data are available for substances in dilute solutions of inert, nonpolar solvents such as hexane (or even better, for substances in the vapor state) will it be possible to make a clear distinction between the effects of changing conformation and changing solvent effects on the optical activity. When such data become available, however, it should be possible to learn a great deal about the little-known field of molecular conformations in general.

3. The optical activity of crystals

Molecules can probably usually pack best into a crystal when they all have the same conformations or at most only a very few different conformations; therefore, in the crystalline state there is usually a high degree of restriction of the amount of freedom of rotation and of orientation of groups about bonds. Aside from this, however, two other effects occur which tend to make the rotations of optically active molecules in crystals different from those of the same molecules in solution. First, there is the occurrence of a "structural" contribution to the optical activity due to the vicinal actions between groups in different molecules. This structural effect is what is responsible for the rotations of crystals made up of the molecules which are themselves not optically active. Second, there is the fact that experimentally the rotation can usually only be measured along the optic axes of the crystal. This prevents the averaging

[5] Rule, J. Chem. Soc. p. 138 (1937).

TABLE II. *Rotations of substances in the solid and liquid states. (Data from I. C. T.).*

SUBSTANCE	ROTATION BY CRYSTAL (IN DEGREES PER DM., Na D LIGHT)	ROTATION IN LIQUID STATE ($[\alpha]_D$)
Open Chain Compounds		
$NaNH_4$ tartrate·$4H_2O$	155	33 in water
NaK " "	76#	33 " "
Rb_2 "	1040	24 " "
Cs_2 "	1700	5 " "
Tartaric acid	609#	15 " "
Zinc malate·$2H_2O$	310#	55 amorphous
$(NH_4)_2$ tartrate	880	29 in water
Asparagine	590	5 " "
Ring Compounds		
$(MoO_4)_2$·$(NH_4)_2$ malate·$2H_2O$	3080	220 for $Na_2(MoO_3)_2C_4H_4O_4$ in H_2O
" Ba " "	3168	
$K_3Ir(C_2O_4)_3$·H_2O	1200	86 in water
$K_3Rh(C_2O_4)_3$·H_2O	640	16 " "
Methyl glucoside (alpha)	440	139 " "
Common camphor	66#	65 molten
Matico "	198	28 "
Patchouli "	133	120 in EtOH
Benzylidene "	818	451 in chloroform
Anisal "	3620	467 in toluene
Corydine	1320	204 in chloroform
Hydrocinchonine sulphate	1280	164 in EtOH
Strychnine	1000	34 in water
Cinchonine antimonyl tartrate	980#	414 molten
Apocinchonine succinate	350	212 in HCl
Quercitol ("strong" axis)	460	26 in water
("weak" axis)	370	
Sucrose ($\lambda = 5790$A)		
("strong" axis)	336#	69 in water ($\lambda = 5790$A)
("weak" axis)	101#	
Substances Active only as Crystals		
Diacetyl phenolphthalein	1980	
Guanidine carbonate	1460	
Ethylene diamine sulphate	1550	
Ammonium oxalate	1200 ($\lambda = 5790$A)	
$LiKSO_4$	350	
HIO_3	5050 ($\lambda = 5790$A)	
SiO_2	2173	
$Pb(CHOO)_2$	1560 ($\lambda = 5790$A)	
$ZnSO_4$·$7H_2O$	241 ($\lambda = 5790$A)	
$NaClO_3$ (Isotropic crystal)	316	
$NaBrO_3$ (" ")	213	
$NaUO_2(C_2H_3O_2)_3$ (" ")	148	
$K_2S_2O_6$	820	

=corrected for density.

of the quantity $\beta \cdot \dot{H}$ over all orientations of the molecule in space in the derivation of the Rosenfeld equation[6] for the optical rotation in terms of the electronic states of a molecule. The usual form of the Rosenfeld equation is then no longer applicable.

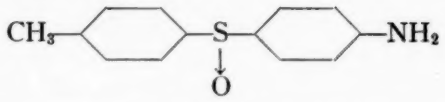
In Table II are given the optical rotations of a number of optically active substances in the solid and liquid (or dissolved) states. The optical rotations of some crystals made up of inactive molecules are also given in order that some idea might be had of the magnitude of the "structural" contribution to the optical activity. The remarkably large rotations shown by most crystals, regardless of the natures of the molecules of which they are made up, is indicative of the operation of many large first-order vicinal actions arising from the severity of the restrictions imposed by crystallization on the freedom of internal and external molecular orientation.

The structural contribution to the optical rotations of crystals is discussed further in connection with solvent effects. below.

[6] E. U. Condon, Rev. Mod. Phys. 9, 432 (1932).

4. Quadrivalent sulphur

The compound



has a very large optical rotation ($[M]_D = 224°$).[7] From the considerations given above and from the probable effective axial symmetries of NH_2 and CH_3, such a large rotation is not possible if only the direct interaction of NH_2, CH_3 and S→O is responsible for the rotation. It seems likely that some force is operating which causes the interactions involving the two benzene rings to be considerably different.

5. Other compounds containing sulphur

The following are the rotations of some sulphur-containing organic acids:[8]

cystein:	$HS\cdot CH_2\cdot CHNH_2\cdot COOH$	$[M]_{5461}^{26°C} = 12\frac{1}{2}°$;	$d\lvert[M]\rvert/dT = -0.12°$ per deg. C
	$HS\cdot CHCH_3\cdot COOH$	$[M]_D^{15°C} = 48°$	
cystine:	$(\cdot S\cdot CH_2\cdot CHNH_2\cdot COOH)_2$	$\frac{1}{2}[M]_{5461}^{26°C} = 386°$;	$\frac{1}{2}d\lvert[M]\rvert/dT = -1.8°$ per deg. C
	$(\cdot S\cdot CH_2\cdot CHCH_3\cdot COOH)_2$	$\frac{1}{2}[M]_D = 510°$;	
	$S(CH_2\cdot CHCH_3\cdot COOH)_2$	$\frac{1}{2}[M]_D = 196°$	
	$(\cdot S\cdot CH_2\cdot CHOH\cdot COOH)_2$	$\frac{1}{2}[M]_D = 13°$.	

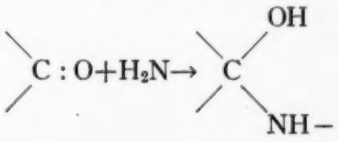
The very large rotations shown by most of those substances in which the sulphur atom is not directly bonded to a hydrogen atom suggest that in these compounds there operate forces which very severely restrict the freedom or orientation of groups. On the other hand, the very large temperature coefficient of the rotation of cystine as compared with that of cystein indicates that these forces are rapidly becoming less important in cystine as the temperature increases. It would be most interesting to know what is the nature of these restricting forces in cystine and the other compounds with large rotations. It would also be interesting to know why they apparently do not operate in $(S\cdot CH_2\cdot CHOH\cdot COOH)_2$.

6. Polypeptides

The following are the rotations of a series of polypeptides of alanine:[9]

	$[M]_D$ IN 2N HCl	$[M]_D \div$ No. OF ALANINE RESIDUE
1-alanine, $C_3H_7O_2N$	+ 13	+ 13
1-alanyl 1-alanine, $C_6H_{12}O_3N_2$	− 58	− 29
di-alanyl 1-alanine, $C_9H_{17}O_4N_3\cdot\frac{1}{2}H_2O$	−186	− 63
tri-alanyl 1-alanine, $C_{12}H_{22}O_5N_4\cdot\frac{1}{2}H_2O$	−358	− 90
tetra-alanyl 1-alanine, $C_{15}H_{27}O_6N_5\cdot H_2O$	−518	−104

It is apparent that as the number of alanine residues in the polypeptide increases, not only does the rotation of the molecule as a whole increase, but there is a marked increase in the rotation per alanine residue. This means that there must be an increasing amount of restriction of the orientation of the groups within the molecule as the polypeptide chain becomes longer. This restriction may arise from the formation of rings involving hydrogen bonds or some other type of bond between amide links, such as[10]:



[7] Harrison, Kenyon and Phillips, J. Chem. Soc. p. 2079 (1926).

[8] Toennies and Bennett, J. Biol. Chem. **112**, 497 (1935). Toennies, Lavine, and Bennett, J. Biol. Chem. **112**, 493 (1935). Louven, J. prakt. Chem. **78**, 63 (1908). Neuberg and Ascher, Biochem. Zeits. **5**, 451 (1907).

[9] Abderhalden and Gohdes, Ber. **64**, 2667 (1931). Levene and Yang, J. Biol. Chem. **99**, 405 (1933).

[10] Wrinch, Phil. Mag. (7) **24** (suppl.), 940 (1937). Mirsky and Pauling, Proc. Nat. Acad. Sci. U.S.A. **22**, 439 (1936).

In the light of the very low rotations shown by the polyhydroxy alcohols, in which purely steric effects would certainly be no less important than they are here, it is difficult to understand how anything not involving bond formation of this type could have such a great effect on the optical rotation. An investigation of the effect of temperature on these rotations would be of interest here.

7. Proteins

Associated with the behavior of the rotatory powers of the polypeptides is that of the proteins. It is known that in the denaturation of egg albumin, serum albumin, casein and edestin,[11] the change in rotation which occurs is not at all marked—the change in the optical rotation during denaturation being of the same order of magnitude as the variation caused by changing the *p*H. This would indicate that the changes in conformation which occur on denaturation are rather small in the immediate neighborhood of any given asymmetric center. There may thus be an analogy between the melting of a solid and the denaturation of a protein: in melting, it is well known that although the long range order typical of the crystal vanishes, the short range order typical of the crystal persists (the number and arrangement of the nearest and next nearest neighbors to a given atom being about the same before and after melting). Denaturation may similarly involve merely a change in the spatial relationships of two widely separated parts of the protein molecule without affecting very markedly the arrangement of atoms in a more restricted region of the molecule. Such a change might result from the breaking of a very few strategically located primary valencies and would not drastically alter the optical activity.

Crystalline proteins in aqueous solutions appear to possess rotations of about the same order of magnitude as other proteins. Thus, pepsinogen has an $[\alpha]_D^{25}$ of $-61°$,[12] pepsin, one of $-70°$,[12] and trypsin, one of about $-50°$.[13] Since the molecular weights of most amino acids average around 100 to 150, these rotations may be taken as a rough measure of the molecular rotations per amino acid residue for these proteins. These rotations are about the same as those found for the noncrystalline proteins mentioned above in connection with denaturation. This similarity between the two is an indication that there is no appreciable difference in the amount of rigidity and of freedom of orientation about asymmetric centers for crystallizable and noncrystallizable proteins; the difference probably arises again from the presence of a higher degree of long range order in the crystallizable proteins.

We may point out the apparently general absence of such large rotations among the proteins as are found in the crystalline solids—a fact which seems to demand a certain looseness in the protein structure, at least when the proteins are in solution. The very large optical rotation of dry gelatin, however (see below), makes a further study of this point desirable.

It has been shown by Smith[14] that there is a definite relationship between the optical activity of a solution of gelatin and other of the properties of the solution, especially its ability to gel. In cooling a solution of gelatin from 35°C to 15°C, there is a reversible increase in the rotatory power from $[\alpha]_D^{35}=-120°$ to $[\alpha]_D^{15}=-270°$. This was proven to arise from a shifting equilibrium between two moles of a high temperature form and one mole of a low temperature form:

$$2A(\text{high temp.})\rightarrow B(\text{low temp.}).$$

Gelatin solutions also metarotate in this temperature region, the kinetics of the reaction corresponding again to two molecules of a high temperature form going to one of a low temperature form. Addition of salts to a solution of gelatin below 15°C causes a change to the high

TABLE III. *Equilibrium in gelatin solutions.*

t°C	$1000/T$	K = EQUILIBRIUM CONSTANT	$-R \ln K$
24	3.3670	0.57	1.12
25	3.3557	1.1	−0.19
26	3.3445	2.0	−1.39
27	3.3333	4.0	−2.77

[11] Pauli and Kölbe, Kolloid Chem. Beihefte **41**, 418 (1935). Pauli and Hoffman, *ibid.* **42**, 34 (1935). Barker, J. Biol. Chem. **103**, 1 (1933).

[12] Northrop, *Crystalline Enzymes* (Columbia University Press, New York, 1939).

[13] Northrop and Kunitz, J. Gen. Physiol. **16**, 267 (1932). The specific rotation of trypsin was calculated from the rotation per milligram of nitrogen per cc ($[\alpha]_D=-0.33°$ per mg N per ml in a one dm tube) and from the elementary analysis of the protein (15 percent N).

[14] Smith, J. Am. Chem. Soc. **41**, 135 (1919).

temperature form to take place. Below 15°C and above 35°C the rotations of gelatin solutions change only very slowly with the temperature. The amorphous solid obtained on evaporating a solution of gelatin at a temperature above 35°C has about the same rotation as the gelatin in the solution from which it was obtained ($[\alpha]_D \cong -100°$), while the rotation of the amorphous solid from a solution evaporated down at a temperature below 15°C has a rotation of $[\alpha]_D = -600$ (assuming a density of 1.27 for solid gelatin and taking Smith's value of $-75°$ per centimeter layer for the rotation).

Analysis (see Table III and Fig. 6) of the temperature dependence of the equilibrium constant found by Smith for these two forms of gelatin gives for the heat of the reaction $2A \rightarrow B$, $\Delta H = -110{,}000$ calories. In order that the equilibrium constant be about equal to unity at 25°C, the entropy in going from $2A$ to B must decrease by about 370 entropy units. That the imposition of such a severe amount of restriction as is implied by an entropy change of 370 entropy units should have no more effect on the rotatory power than to change it from $-120°$ to $-270°$ is perhaps surprising in the light of the fact that evaporation of a cold gelatin solution can bring about a further increase in the levorotation to $-600°$. This indicates that the removal of water from gelatin can probably impose more restrictions on the movements of the groups in the gelatin molecule than is observed to occur in the change from the high temperature form of the protein to the low temperature dimer. The behavior of gelatin solutions is understandable if, as was suggested for denaturation (where similar exceptionally large heat and entropy effects are known to be involved[15]) the changes which occur involve chiefly an alteration in the long range order of the molecules, accompanied by moderate changes in the arrangement of the atoms in the immediate neighborhood of any given atom. On evaporation of a cold gelatin solution, this tightening of the arrangement about the asymmetric centers appears to be rather more marked, even approaching the state of affairs in crystalline solids.

In the study of such a complex problem as that of the structure of the proteins, it is of more than ordinary importance that all possible methods of attack on the problem be utilized. The optical activities of the proteins are capable of being applied as such a tool and are more sensitive than most properties to variations in experimental conditions. It is unfortunate that they are not being studied more intensively than they are, since they would appear to be capable of answering many structural questions.

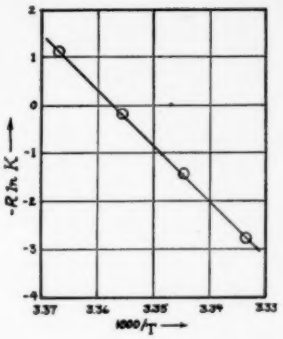


FIG. 6.

8. Solvent effects on the optical rotatory power

A liquid or solid may be considered to be a large molecule made up of many smaller molecules between which there operate forces of attraction and repulsion of some sort. The very fact that one substance can dissolve in another can only be a result of the forces which act between the two different kinds of molecules. We may say, therefore, that when one substance is dissolved in another, a compound is formed between the molecules involved. The properties of this compound may bear a very complex relationship to those of the molecules of which it is composed, and in general it is the understanding of this relationship which is the object of the study of what are usually spoken of as solvent effects.

From this point of view, the effects of solvents on the rotatory power are merely the result of the formation of some kinds of compounds between the solvents and the optically active molecules concerned, and in calculating completely the optical rotation of a solution (or of any liquid or solid, for that matter), we must consider not only the optically active molecule itself, but its environment as well. This environment may affect the optical rotation by changing the vicinal actions within the active molecule (e.g., through a distortion of the molecular

[15] Eyring and Stearn, Chem. Rev. **24**, 253 (1939).

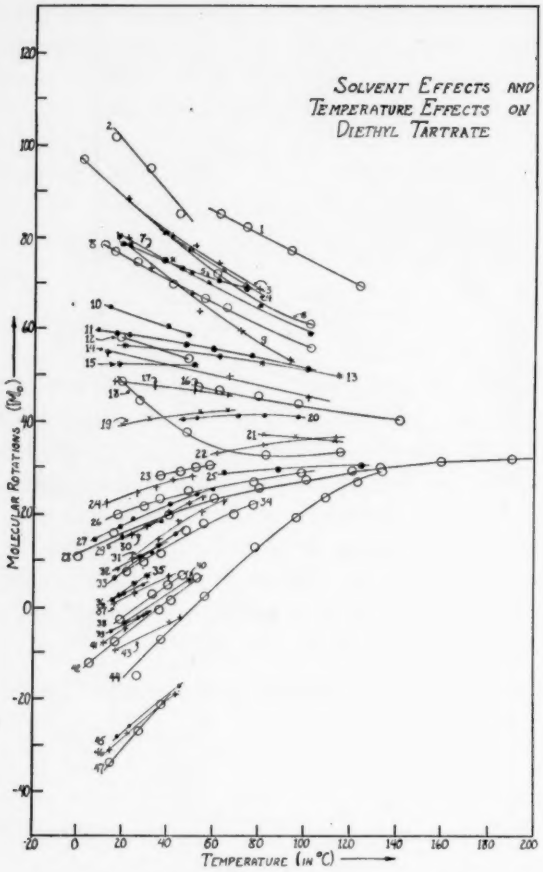


FIG. 7. Solvent effects and temperature effects on diethyl tartrate.

KEY TO FIG. 7

1. alpha-nitro naphthalene
2. meta toluidine
3. benzaldehyde
4. paratoluidine
5. aniline
6. pyridine
7. nitrobenzene
8. orthonitro toluene
9. phenol
10. orthotoluidine
11. meta nitro toluene
12. benzyl alcohol
13. salicyl aldehyde
14. water
15. ortho $CH_3 \cdot O \cdot C_6H_4 \cdot NO_2$
16. 2,6, dinitro toluene
17. methyl aniline
18. quinoline
19. ortho $C_2H_5 \cdot O \cdot C_6H_4 \cdot NO_2$
20. paranitro toluene
21. meta dinitro benzene
22. orthonitro phenol
23. para $CH_3 \cdot O \cdot C_6H_4 \cdot NO_2$
24. methanol
25. 2,4, dinitro toluene
26. glycerol
27. ethanol
28. no solvent
29. $C_6H_5 \cdot OC_2H_5$
30. *n*-propanol
31. iso-butanol
32. anisol
33. dimethyl aniline
34. secondary octanol
35. carbon tetrachloride
36. methyl iodide
37. bromoform
38. methylene iodide
39. methylene chloride
40. ethyl iodide
41. ethylidene chloride
42. chloroform
43. ethylene chloride
44. anti-benzaldoxime
45. acetylene tetrabromide
46. ethylene bromide
47. acetylene tetrachloride

framework), or it may affect the optical activity by itself exerting vicinal actions on the chromophores of the active molecules and by the vicinal actions of the groups in the active molecule on the chromophoric groups in the surrounding solvent molecules.

The effect of freedom of orientation and rotation of groups on the rotatory power has an important bearing when we come to consider the magnitude of the solvent effect. The distances between molecules in solutions are considerably greater than those between groups on an asymmetric carbon atom; therefore, in order to be appreciable, any vicinal actions which operate between groups in two different molecules must be first-order vicinal actions. But these will obviously vanish unless the two molecules are kept from taking on all possible orientations with respect to one another; that is, in order that there can be large vicinal actions between two molecules, the forces acting between them must tend to cause a considerable amount of orientation of the molecules with respect to one another. (Note, however, that even when neighboring molecules can take up all possible orientations with respect to one another, there may still be a nonvanishing though small contribution to the rotation due to higher order vicinal actions.) Thus, forces such as van der Waals interactions, which depend predominantly on intermolecular distance rather than on the relative orientations of molecules, should result in only negligible vicinal actions between solvent and solute molecules. This probably accounts for the small solvent effects which have been found for the saturated hydrocarbon pinane, whose rotation is nearly constant in most solvents, once secondary effects of the refractive index have been taken into account.[16]

We may say, then, that those factors which reduce the directional character of the interaction between solvent and solute will generally reduce the solvent effect on the rotatory power. Now, one of the most effective ways of reducing a relatively large frequency of occurrence of certain preferred orientations of solvent and solute molecules is to increase the temperature. That is, by increasing the temperature, solvent effects should decrease, and when the temperature is raised, the rotations of a given substance in different solvents should tend to approach one another (correction having been made for the

[16] Rule and Chambers, J. Chem. Soc. p. 145 (1937).

different refractive indices of different solvents by dividing the rotation[2] by $(n^2+2)/3$. This has frequently been noticed,[17] but the effect is most strikingly shown by the very completely investigated case of diethyl tartrate. In Fig. 7 the temperature dependence of the rotations of diethyl tartrate in various solvents is shown. It will be seen that at high temperatures the rotations of diethyl tartrate in all solvents tend to have a value of about 35°–45°. If the rotations used in the figure had been corrected for the effect of the refractive index, the tendency of all of the curves to converge on a single rotation would presumably have been even more marked.[18]

The so-called "structural" contribution to the optical rotatory power of solids (that is, the difference between the rotatory power of a crystal averaged over all orientations of the crystal with respect to the light wave, and that of the same substance in the dilute vapor) is really only another aspect of a solvent effect, this time carried to an extreme. It is again merely due to interactions which occur between neighboring molecules which have taken up very definite orientations with respect to one another, so forming a clearly defined "solvent-solute" complex. In the case of those crystals, such as quartz, or better, diacetyl phenolphthalein, which are optically active in spite of the fact that the molecules of which the crystal is made are themselves inactive, we thus arrive at the interesting conclusion that all of the optical rotatory power arises from "solvent effects" of one kind or another.

An interesting exception to the tendency of the structural contributions to be large in solids is to be found in camphor, where the optical rotations of the solid and liquid forms are very nearly the same. This is doubtless due to the fact that camphor molecules are nearly freely rotating with respect to one another in the crystal as well as in the melt—a fact which is known from the study of the dielectric constant of solid camphor.[19]

[17] Lowry, *Optical Rotatory Power* (Longmans, Green and Co., New York, 1935), p. 353.

[18] The data used in preparing Fig. 7 were obtained from a series of papers by Patterson and co-workers in J. Chem. Soc. London, 1901 ff.

[19] White and Morgan, J. Am. Chem. Soc. **57**, 2078 (1935).

9.

A number of optically active substances are known in which the axes of symmetry of the groups attached to the asymmetric carbon atom all intersect in one point. Where the groups are linear or the axes are threefold, the optical rotations which are observed must be due to vicinal actions of higher order than the first. The following are the rotations of some of these substances:

$$\begin{matrix} & H & \\ CH_3\cdot & C & \cdot CN \\ & Br & \end{matrix} \qquad [M]_D = 21°\ ^{20}$$

$$\begin{matrix} & H & \\ CH_3\cdot & C & \cdot CN \\ & NH_3^+ & \end{matrix} \qquad [M]_D = 13°\ ^{21}$$

$$\begin{matrix} & I & \\ H\cdot & C & \cdot SO_3^- \\ & Cl & \end{matrix} \qquad [M]_D = 36°\ ^{22}$$

The orders of magnitude of the rotations which are found here should be taken as typical of those to be expected for class *C* aliphatic compounds. It would be very interesting to have data on the effects of temperature and solvent on the rotations of these compounds in order to evaluate the importance of first-order interactions with neighboring solvent molecules in determining the value of the optical rotatory power.

10.

Whether or not the optical rotatory power of a given molecule arises from first-order vicinal actions is a question of great importance when one is trying to calculate the optical rotation. It is well to point out the probable unsoundness of any theory with attempts to explain optical activity without taking into account first-order vicinal actions. Until a more complete knowledge of these is available, it is hard to see how progress can be made in understanding higher order vicinal actions, since these are probably best regarded as a result of the compounding of first-order effects.

[20] Berry and Sturtevant, J. Am. Chem. Soc. **61**, 3583 (1939).

[21] Delepine, Bull. Soc. Chim. (3) **29**, 1195 (1903).

[22] Read and McMath, J. Chem. Soc. p. 273 (1932).

The Electron Diffraction Investigation of the Molecular Structures of Tetrahydrofuran, 2,5-Dihydrofuran, Furan, Dioxadiene and *p*-Dioxene*

J. Y. BEACH
Princeton University, Princeton, New Jersey
(Received October 28, 1940)

THE molecular structures of the above compounds have been determined by electron diffraction. The compounds were all prepared by Professor R. K. Summerbell of Northwestern University. The interpretation of the photographs was carried out by the radial distribution method[1] and by the visual correlation method.[2] In calculating theoretical intensity curves atomic numbers were used in place of atomic form factors except in the case of hydrogen in which case the value 1.2 was used for the atomic number.[3] Temperature factors were used for terms involving hydrogen atoms.

Tetrahydrofuran.—The electron diffraction photographs of tetrahydrofuran, taken with the sample at 22°C, show eight maxima and minima to s equals 22. The radial distribution function calculated from these data shows two principal peaks. The first is at 1.49A and the second at 2.42A. The first of these is an average of the C−O single-bond distance and the C−C single-bond distance, and the second an average of the distances across the ring.

Theoretical intensity curves were calculated for eleven models. Only planar models were considered. The oxygen bond angle was varied from 90° to 130°. The ratio of C−O distance to C−C distance was varied from 1.47/1.54 to 1.39/1.54. The values from the table of covalent radii[4] are 1.43 and 1.54A. The final results are: oxygen bond angle = 111±2°, C−O = 1.43 ±0.03A and C−C = 1.54±0.02A.

2,5-Dihydrofuran.—The electron diffraction photographs, taken with the sample at 25°C, show seven maxima and minima to s equals 19. The two principal peaks on the radial distribution function are at 1.45 and 2.36A. The first is an average of the C−O distance, the C−C distance and the C=C distance. The second is an average of the distances across the ring.

Theoretical intensity curves were calculated for thirteen planar models. The oxygen bond angle was varied from 90° to 130°. The ratio of C−O distance to C−C distance was varied from 1.43/1.58 to 1.43/1.46. The C=C distance was assumed to be 1.34A. The values from the table of covalent radii[4] for the C−O, C−C and C=C distances are 1.43, 1.54 and 1.34A, respectively. The final results are: oxygen bond angle = 110 ±3°, C−O = 1.43±0.03A. C−C = 1.54±0.03A and C=C = 1.34 (assumed).

Furan.—Photographs taken with the sample at 0°C show eight maxima and minima to s equals 23. The results we obtain are in agreement with those published for this compound by Schomaker and Pauling.[5] The radial distribution function shows peaks at 1.39 and 2.27A. The final results are: oxygen bond angle = 107±3°, C−O = 1.40±0.03A, C−C = 1.46±0.03A and C=C = 1.35 (assumed).

Dioxadiene.—Photographs taken with the sample at 35°C show seven maxima and minima to s equals 19. The radial distribution function shows two principal peaks. The first at 1.40A is an average of the C−O distance and the C=C distance. The second at 2.40A is an average of the distances across the ring.

Theoretical intensity curves were calculated for nine planar models. The oxygen bond angle was varied from 110° to 120°. The ratio of C−O distance to C=C distance was varied from 1.47/1.34 to 1.36/1.36. The final results are: oxygen bond angle = 116±4°, C−O = 1.41 ±0.03A and C=C = 1.35±0.03A.

Dioxene.—Electron diffraction photographs

* Presented at the Symposium on the Structure of Molecules and Aggregates of Molecules at the Fifth Annual Symposium of the Division of Physical and Inorganic Chemistry of the American Chemical Society, Columbia University, New York, December 30, 1940 to January 1, 1941.

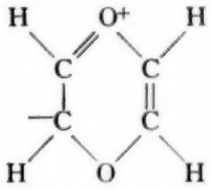
[1] Walter and Beach, J. Chem. Phys. **8**, 601 (1940).
[2] Pauling and Brockway, J. Chem. Phys. **2**, 867 (1934).
[3] Bauer, J. Am. Chem. Soc. **60**, 524 (1938), see footnote p. 526.
[4] Pauling and Huggins, Zeits. f. Krist. **87**, 205 (1934).
[5] Schomaker and Pauling, J. Am. Chem. Soc. **61**, 1769 (1939).

taken with the sample at 45°C show seven maxima and minima to s equals 19. The interpretation of the photographs has not been completed at the time of writing.

Discussion

The interatomic distances found in tetrahydrofuran and 2,5-dihydrofuran are just those predicted by the table of covalent radii.[4] In furan the shortening of the C−C bond and the C−O bond observed by Schomaker and Pauling[5] has been confirmed.

The C−O distance in dioxadiene is slightly less than the sum of the single covalent radii of oxygen and carbon. This is presumably due to some contribution of structures of the type



to the electronic structure of the molecule. A correlation of the molecular structures of *p*-dioxane, *p*-dioxene and dioxadiene with their solubilities in water will be given.

This investigation will be published jointly with Professor R. K. Summerbell who prepared and purified the compounds and measured the solubilities of the dioxenes.

JANUARY, 1941 JOURNAL OF CHEMICAL PHYSICS VOLUME 9

The Influence of Intramolecular Atomic Motion on Electron Diffraction Diagrams*

P. Debye
Department of Chemistry, Cornell University, Ithaca, New York
(Received October 19, 1940)

A discussion of the influence of atomic vibrations and of free or hindered rotations of molecular groups in molecules on their scattering properties. The radial distribution curve is calculated and the results which may be expected from an analysis of this curve are indicated.

SINCE the electron-diffraction method has been improved by the use of the rotating sector,[1] it is to be expected that in the near future more detailed information about the intensity distribution of the scattered electrons will be available. It has already been indicated that the diffraction pictures may then serve not only to determine the average values of atomic distances but should also give information as to their natural degree of indeterminacy. In this paper the question will be discussed from the theoretical point of view.

Suppose a free molecule consisting of atoms $1\cdots n\cdots$ is struck by a primary electron beam of intensity J_0, then at a large distance R (large compared with molecular dimensions) a secondary intensity J will be observed, which for an appropriate range of electron speeds can be represented by the formula

$$\frac{J}{J_0}=\frac{(\epsilon^2/E)^2}{R^2}\frac{1}{s^4}\{\sum_n g_n+\sum_n (z_n-f_n)^2 + \sum_{m,n}{}'(z_m-f_m)(z_n-f_n)A_{mn}\}. \quad (1)$$

The letter ϵ indicates the electronic charge and E is the kinetic energy of an electron in the primary beam; the quotient ϵ^2/E, therefore, has the dimension of a length which for 50-kv electrons is equal to $2.88\cdot 10^{-12}$ cm. The second factor $1/s^4$ in which $s=2\sin(\theta/2)$ (θ being the angle between the primary and the secondary beams) represents Rutherford's well-known scattering factor. The letter z_n represents the number of

* Presented at the Symposium on the Structure of Molecules and Aggregates of Molecules at the Fifth Annual Symposium of the Division of Physical and Inorganic Chemistry of the American Chemical Society, Columbia University, New York, December 30, 1940 to January 1, 1941.

[1] P. P. Debye, Physik. Zeits. 40, 66 and 404 (1939).

elementary charges contained in the nucleus of atom n, whereas f_n and g_n refer to the two scattering factors of the electronic shell, the first, f, measuring the coherent (elastic) and the second, g, the incoherent (inelastic) scattering. If the electron wave-length is indicated by λ both factors are functions of the quotient s/λ and can be taken from well-known tables, which are being used for x-ray scattering purposes. The scattering factor g_n begins with the value 0 for $s=0$ and ultimately rises to z_n; the factor f_n for the coherent scattering begins with the value z_n and approaches 0 for large values of s. Together with the factor $1/s^4$ the two first sums in Eq. (1), extending over all the atoms contained in the molecule, describe the intensity distribution of a background, which shows no maxima and minima. The third sum extends over all the connecting lines between any two different nuclei m and n and contains the interference functions A_{mn}. Such a function is defined as the average value of $\exp(ik\bar{s}\bar{r}_{mn})$ where k equals $2\pi/\lambda$, $\bar{s}$ is a vector of length unity connecting the ends of two vectors of length unity issuing from one common origin, the one in the direction of the primary and the other in the direction of the secondary beam, whereas $\bar{r}_{mn}$ is a vector connecting the two nuclei m and n. If the molecule is absolutely rigid, this average value is found to be equal to $\sin ksr_{mn}/ksr_{mn}$, and this is the expression for A_{mn} which has been in common use to determine, with the help of the third sum in Eq. (1), the position of the interference maxima. In this paper we are concerned with the changes in A_{mn} due to the relative motion of the nuclei.

The simplest case which can be considered is that of a diatomic molecule with two like atoms. In this case only one scattering function A_{12} occurs in Eq. (1). Keeping the center of gravity of the molecule fixed in space and calling the displacement of one of the atoms in the direction of the bond ξ, the atomic distance will be equal to $l+2\xi$ and the probability that ξ has a value between ξ and $\xi+d\xi$ equals $\exp(-\xi^2/a^2)/a\sqrt{\pi}$. This probability function has been shown by F. Bloch to be valid for all temperatures, provided the constant a is so adjusted that the average energy of the atomic vibration of frequency ν becomes equal to Planck's well-known expression $u=h\nu/2+h\nu/(\exp[h\nu/\beta T]-1)$. In our particular case this condition leads to the value

$$a^2=\frac{1}{4\pi^2 m\nu^2}\left[\frac{h\nu}{2}+\frac{h\nu}{e^{h\nu/\beta T}-1}\right], \qquad (2)$$

in which m denotes the mass of one of the atoms. It is to be remarked that in many cases at ordinary temperatures $h\nu$ proves to be large compared with βT. The indeterminacy of position is then only due to the zero-point energy $h\nu/2$ and generally amounts to a few percent of the average atomic distance.

In order to calculate A_{12} it has to be borne in mind that not only does the distance $r_{12}=l+2\xi$ show fluctuations, but that during the rotation of the molecule the direction of $\bar{r}_{12}$ may make any angle with the fixed direction of $\bar{s}$. This leads to the equation

$$A_{12}=\int_{-\infty}^{+\infty}\int_0^{\pi} e^{iks(l+2\xi)\cos\vartheta}\frac{\sin\vartheta}{2}d\vartheta e^{-\xi^2/a^2}\frac{d\xi}{a\sqrt{\pi}} \qquad (3)$$

if the angle between $\bar{r}_{12}$ and $\bar{s}$ is indicated by ϑ. Performing first the integration with respect to ξ and substituting v for $\cos\vartheta$, Eq. (3) reduces to

$$A_{12}=\frac{1}{ksl}\int_0^{ksl} e^{-(a^2/l^2)v^2}\cos v\,dv. \qquad (3')$$

Bearing in mind that a/l is a small quantity, it can be shown that, neglecting parts of the order a^2/l^2 or less, A_{12} can be expressed by the formula

$$A_{12}=\frac{e^{-k^2s^2a^2}}{[1+4(a^2/l^2)k^2s^2a^2]^{\frac{1}{4}}}\frac{\sin(ksl-\varphi)}{ksl}, \qquad (4)$$

where the angle φ is defined by the equation

$$\operatorname{tg}\varphi=2(a/l)ksa. \qquad (4')$$

The most important part of the interference function A_{12} is $\exp(-k^2s^2a^2)$, for although a is small compared with l the exponent $k^2s^2a^2$ will obtain values of the order unity for scattering angles θ, corresponding to interference maxima of a rather low order. The 5th maximum of $\sin x/x$ occurs for a value $x=ksl=10\pi+\pi/2=33$. If a equals 3 percent of l, the value of ksa at the same angle will be $(a/l)ksl=1$ and the amplitude of the interference factor A_{12} will be reduced to 0.37 times the value it would have if no fluctua-

tions of the atomic distance existed. Equation (4) reveals further that a shift in the position of the maxima of higher order is to be expected, which according to (4′) increases with increasing scattering angle. However, in our example the influence of the correcting angle φ would change the value of x by only $\frac{1}{5}$ of one percent. Finally, the effect of the square root in the denominator of Eq. (4) is, in the region of θ here considered, still less important. As a result of this discussion, we can, therefore, conclude that the equation

$$A_{12}=e^{-k^2s^2a^2}\sin ksl/ksl \qquad (5)$$

will be an approximation sufficient for all the scattering angles, which are of any importance practically. At the same time it is to be noticed that the actual measurement of the amount of indeterminacy of position of the atoms by observing the decrease of intensity of the interference rings with increasing order seems to be well within the range of possibilities. From the form of the probability function for ξ it follows that the average of ξ^2 equals $a^2/2$. If at any moment the distance of the two atoms is put equal to $l+\Delta$, the deviation Δ of the average distance l is 2ξ giving for the average value of the square of Δ the value $4\xi^2=2a^2$. All we have to do in order to take account of the indeterminacy of the atomic positions is, according to Eq. (5), to introduce the correction factor $\exp[-(k^2s^2/2)\Delta^2_{Av}]$. It can be shown that this is not only true for a diatomic molecule, but for any molecule, every interference factor A_{mn} of our general formula Eq. (1) being, of course, corrected by a correction factor corresponding to the fluctuation Δ_{mn} of the distance l_{mn} in question.[2] The theoretical calculation of Δ^2_{Av}, however, is a simple procedure only for a diatomic molecule. In this case we can apply Eq. (2), but for a more complicated molecule the vibration of its frame has first to be decomposed into free vibrations, and it is necessary not only to know the vibration frequencies but also to determine the contribution of every such free vibration to the actual displacement of every atom. As an example, we may refer to the case of the CO_2 molecule. As is well known, for this linear molecule three kinds of free vibration exist, two in which the displacement of the atoms is in the direction of the line connecting the atoms, and the third perpendicular to this line. This last free vibration does not contribute to Δ^2_{Av}, at least according to the theory in existence, which is only a first approximation for small displacements. If ν_1 is the smaller, and ν_2 the higher, frequency of the first two vibrations, α the elastic constant of the bond between the atoms C and O and if we denote by u_1 and u_2 the energy contents of two linear resonators of frequencies ν_1 and ν_2 (according to Planck's formula including the zero-point energy), it is found that for the C−O bond

$$\Delta_{Av}{}^2=\frac{1}{2\alpha}(u_1+u_2)$$

and for the O−O bond

$$\Delta_{Av}{}^2=\frac{1}{2\alpha}u_1.$$

With $\nu_1=4.10^{13}$ and $\nu_2=7.10^{13}$ for the CO_2 molecule the average amplitude of the fluctuations of the C−O distance (defined as the square root of Δ^2_{Av}) is 0.035A and that of the O−O distance is 0.021A. The whole indeterminacy of position is practically due to the zero-point energy and should, therefore, show only a very slight increase with increasing temperature.

Another kind of relative motion of the atoms exists in molecules with so-called free rotating groups. If we take as an example the molecule CH_2Cl-CH_2Cl the scattering function A related to the Cl−Cl distance is the one here to be considered. The discussion of two extreme cases will illustrate the behavior of the scattering function: Case (a) free rotation; case (b) prevalence of one position.

(a) The scattering function will, in general, be expressed by an integral, which can only be treated by graphical methods. However, in the range of maxima of higher order a very convenient and simple representation can be derived. For the sake of brevity we confine our attention to the molecule CH_2Cl-CH_2Cl, although the method is quite general. In calculating the average, which is going to represen the scattering function A, we can first consider a definite position of the two Cl atoms with respect to each other, and average over all possible orientations of this special kind of

[2] R. W. James, Physik. Zeits. **33**, 737 (1932).

molecule. This leads to the equation

$$A=\int \frac{\sin ksr}{ksr}dW, \tag{6}$$

where dW represents the probability of finding a molecule in the immediate neighborhood of the configuration characterized by the distance r of the two Cl atoms. Taking into account all possible configurations, the distance r has a minimum value l_1, and a maximum value l_2. If such is the case it can be shown that, for large values of ks, integrals of the kind represented by Eq. (6) can be expressed asymptotically by considering only the parts related to the neighborhood of the minimum or maximum value of r. This asymptotic expression generally is a good approximation even for rather small values of the argument.

Call a the radius of the circle in which either of the Cl atoms is supposed to move freely and b the distance of the two planes in which the two Cl atoms rotate. If, further, φ is the angle of rotation related to the motion of one of the Cl atoms in its plane, measured from the position $r=l_1=b$ and ψ a similar angle measured from the position $r=l_2=(b^2+4a^2)^{\frac{1}{2}}$, we have as a first approximation

$$r=l_1+\frac{a^2}{l_1}\frac{\varphi^2}{2}$$

or

$$r=l_2-\frac{a^2}{l_2}\frac{\psi^2}{2}.$$

Inserting these expressions in Eq. (6) and writing for the denominator ksr its minimum or maximum value, it follows that A is the imaginary part of the expression

$$\frac{e^{iksl_1}}{ksl_1}\int_{-\infty}^{+\infty} e^{i(ksa^2/2l_1)\varphi^2}\frac{d\varphi}{2\pi}+\frac{e^{iksl_2}}{ksl_2}\int_{-\infty}^{+\infty} e^{-i(ksa^2/2l_2)\psi^2}\frac{d\psi}{2\pi}.$$

Our asymptotic formula for A, therefore, becomes

$$A=\frac{1}{\sqrt{2\pi}}\frac{\sin(ksl_1+\pi/4)}{ksa(ksl_1)^{\frac{1}{2}}}+\frac{1}{(2\pi)^{\frac{1}{2}}}\frac{\sin(ksl_2-\pi/4)}{ksa(ksl_2)^{\frac{1}{2}}}. \tag{7}$$

In case (a) of free rotation we will have, according to Eq. (7), a superposition of two periodic terms, related to the smallest and the largest distance of the two atoms, whereas at the same time the intensity will decrease as $s^{-\frac{1}{2}}$ and not as s^{-1}, characteristic for the usual case of an invariable atomic distance.

(b) From experiments with x-rays F. Ehrhardt[3] came to the conclusion that in ethylenedichloride we have no free rotation but a decided preference for the *trans*-position. This was confirmed in the electron diffraction experiments of J. Y. Beach and K. J. Palmer,[4] who also estimated the potential energy involved in turning from the *trans*- to the *cis*-position at more than 5 kcal. per mole. If this potential energy is denoted by V the probability function for small values of the angle ψ will have the form

$$\left(\frac{V}{4\pi\beta T}\right)^{\frac{1}{2}}e^{-(V/4\beta T)\psi^2}d\psi.$$

In this case (b) only positions in the neighborhood of the *trans*-position have to be considered, as $V/\beta T$ is large (of the order 10). Now it follows that the interference function A can be expressed as the imaginary part of the integral

$$\frac{e^{iksl_2}}{ksl_2}\left(\frac{V}{4\pi\beta T}\right)^{\frac{1}{2}}\int_{-\infty}^{+\infty} e^{-(V/4\beta T)\psi^2}e^{-i(ksa^2/2l_2)\psi^2}d\psi.$$

In this way we arrive at the formula

$$A=(\cos\varphi)^{\frac{1}{2}}\frac{\sin(ksl_2-\varphi/2)}{ksl_2},\qquad \operatorname{tg}\varphi=\frac{2a}{l_2}\frac{\beta T}{V}ksa, \tag{8}$$

which for values of ksa large compared with $V/\beta T$ (as the angle φ approaches $\pi/2$) is ultimately transformed into

$$A=\left(\frac{V}{2\beta T}\right)^{\frac{1}{2}}\frac{\sin(ksl_2-\pi/4)}{ksa(ksl_2)^{\frac{1}{2}}}, \tag{8'}$$

making the interference function again proportional to $s^{-\frac{1}{2}}$ for large scattering angles. In order

[3] F. Ehrhardt, Physik. Zeits. **33**, 605 (1932).
[4] J. Y. Beach and K. J. Palmer, J. Chem. Phys. **6**, 639 (1938).

to separate the different terms of the last sum in Eq. (1) the so-called "radial distribution method," first invented for the analysis of x-ray diagrams of liquids, has been applied to the electron diffraction photographs of single molecules. The skill of L. Pauling and his co-workers made the method successful for the tracing of atomic distances with the help of peaks in the radial distribution curve, in spite of the fact that the intensities of the interference maxima on which the calculation of this curve have been based were only visual estimates derived from photographs which in reality do not show maxima at all, combined with a rough summation process which replaces the actually prescribed integration. We can ask how much more the radial distribution curve can be expected to reveal if we start with photographs in which the disturbing influence of the steeply sloping background has been experimentally eliminated by the use of the rotating sector.

The interference function A_{mn} in Eq. (1) is represented by the expression $\sin ksl/ksl$ only if the distance l is taken to be invariable. If $P(l)dl$ is the probability that l has a value between l and $l+dl$ the correct form for the interference function (replacing ks by the letter σ) is

$$A=\int P(l)\frac{\sin \sigma l}{\sigma l}dl.$$

The application of the rotating sector has the effect of multiplying the scattered intensities with a factor proportional to the reciprocal of

$$\frac{1}{s^4}\sum_n (z_n-f_n)^2.$$

It is conceivable that some further improvement could be obtained by making the opening of the sector correspond to the reciprocal of

$$\frac{1}{s^4}\left[\sum_n (z_n-f_n)^2+\sum_n g_n\right]$$

in order to include both elastic and inelastic scattering. In both cases the coefficient of A_{mn}, which we will call c_{mn} or c, tends to become a constant for the maxima of higher order. In order to obtain the function of the arbitrary variable r, which is going to represent the radial distribution curve, the ordinates of the underlying intensity curve are multiplied with $\sigma \sin \sigma r$ and an integration over σ is carried out. Every term of the last sum in Eq. (1) is, therefore, going to contribute to the radial distribution curve an amount equal to

$$\begin{aligned} D(r)&=\int_0^\infty cA\sigma \sin \sigma r d\sigma \\ &=\int\int c(\sigma)\frac{P(l)}{l}\sin \sigma l \sin \sigma r d\sigma dl. \end{aligned} \tag{9}$$

If, now, we neglect the variability of the coefficient $c(\sigma)$, which only affects the first maxima, the integration with respect to σ can be carried out. Letting σ go from 0 to a large value σ_0 we obtain in this case

$$D(r)=\frac{c}{2}\int \frac{P(l)}{l}\left[\frac{\sin \sigma_0(r-l)}{(r-l)}-\frac{\sin \sigma_0(r+l)}{(r+l)}\right]dl.$$

In the limit for $\sigma_0=\infty$ only the immediate neighborhood of $l=r$ and $l=-r$ is of any importance, and if we restrict ourselves to positive values of the distance r, our expression for D becomes

$$D(r)=\frac{\pi}{2}c\frac{P(r)}{r}. \tag{10}$$

The ordinates of the distribution curve, therefore, represent essentially the probability of the occurrence of the atomic distances r in the molecule. If, for instance, only vibrations exist in the molecule and one of the distances considered has an indeterminacy characterized by an average square displacement Δ^2 the ordinates of the radial distribution curve in the vicinity of $r=l$ will be represented by the equation

$$D(r)=\frac{\pi}{2}\frac{c}{l}\frac{e^{-(r-l)^2/2\Delta^2}}{\Delta(2\pi)^{\frac{1}{2}}}.$$

The surface integral of this peak is

$$\int D(r)dr=\frac{\pi}{2}\frac{c}{l}$$

and, since c is proportional to the product of the numbers z_m and z_n characterizing the charges of

the nuclei at both ends of the connecting bond, this integral reveals the kind of atoms involved. The better radial distribution curve here considered will presumably be of special importance in the case of molecules containing rotating groups. It ought to give more detailed evidence of the potential energy involved in the restraint opposing free rotation if interpreted according to Eq. (10). As an example we may state that for the case of a molecule like CH_2Cl-CH_2Cl the part of the radial distribution curve related to the Cl−Cl distance would be represented by

$$D(r)=\frac{c}{\pi}\frac{1}{((r^2-l_1^2)(l_2^2-r^2))^{\frac{1}{2}}},$$

assuming absolutely free rotation. In an actual case the infinities for the smallest and the largest distances l_1 and l_2 would not occur. They result from the calculation only because vibrations of the atoms have been neglected.

The foregoing very simple interpretation of the radial distribution function has been made possible by neglecting the variability of the coefficients $c(\sigma)$. This variability is due to the shielding effect of the atomic electrons. Consequently Eq. (10) represents an ideal curve as it would be obtained if the scattering of the molecule were purely nuclear. It remains to be seen how much the ideal curve actually is distorted by the electronic scattering. For exact calculations of this effect a graphical method has to be employed based on tables for f and g. However, a rough estimate as to the importance of the effect can be made in the following way. Let us assume an artificial case in which coefficient c can be represented as a function of σ by the expression

$$c=c_0(1-e^{-a\sigma}). \tag{11}$$

This can be taken as an approximation because it expresses in a general way the characteristic behavior of the coefficient $(z_m-f_m)(z_n-f_n)$ which begins with the value 0 and as σ increases tends to the constant value z_mz_n. The constant a introduced in (11) has the dimension of a length and will be of the order of 1A. If this representation is accepted, the actual radial distribution curve is obtained by subtracting from the function

$$D(r)=\frac{\pi}{2}c_0\frac{P(r)}{r}. \tag{12}$$

values of another function $D'(r)$ which follows from Eq. (9) by substituting $c_0e^{-a\sigma}$ instead of $c(\sigma)$. The integration with respect to σ can be performed with the result

$$D'(r)=\frac{c_0}{2a}\int\left[\frac{1}{1+\left(\frac{r-l}{a}\right)^2}-\frac{1}{1+\left(\frac{r+l}{a}\right)^2}\right]P(l)dl.$$

Suppose, now, that we are dealing with an atomic distance, which, due to vibrations, shows fluctuations characterized by an average square displacement Δ^2, small compared with a^2. Under these circumstances $D'(r)$ is readily seen to be equal to

$$D'(r)=\frac{c_0}{2al}\left[\frac{1}{1+\left(\frac{r-l}{a}\right)^2}-\frac{1}{1+\left(\frac{r+l}{a}\right)^2}\right], \tag{12'}$$

where l now stands for the average atomic distance. The actual radial distribution curve, represented by the function $D-D'$, consists, therefore, of the superposition of two peaks, a sharp positive peak of width Δ and a height proportional to $1/\Delta$ combined with a broad negative peak of width a and a height proportional to $1/a$. From this discussion we conclude that the shielding effect of the atomic electrons introduces corrections of minor importance, which however in the case considered will be responsible for the occurrence of negative secondary peaks in the vicinity of the main positive peak. Given a fairly precise actual radial distribution curve it should not be difficult to transform it satisfactorily into an ideal curve by adding to its ordinates values represented by a function $D'(r)$ calculated with the help of tabulated atomic scattering factors instead of using a more or less adequate interpolation formula as we did here in our discussion. In the case of rotating groups, where not a single sharp peak can be considered, it may even be necessary to use a method of trial and error in order to accomplish the transformation.

Potential Energy Functions for Diatomic Molecules*

HUGH M. HULBURT AND JOSEPH O. HIRSCHFELDER
Chemistry Department, University of Wisconsin, Madison, Wisconsin

(Received October 9, 1940)

The usual Morse functions are determined from the energy of dissociation, the equilibrium separation of the nuclei, and the fundamental vibration frequency. Two additional spectroscopic constants, $\omega_e x_e$ and α_e, are available for most of the common diatomic molecules and permit us to add a two-parameter correction term to the Morse curve. Both the potential $V/D=(1-e^{-x})^2+cx^3(1+bx)e^{-2x}$ and the extended Morse curve of the Coolidge, James and Vernon type, $V/D=C_2(1-e^{-x})^2+C_3(1-e^{-x})^3+C_4(1-e^{-x})^4$ agree with accurate potentials in those cases where they are known. Here $x=2\beta(r-r_e)/r_e$. The constants for the first of these potentials are easy to evaluate and are given for 25 common diatomic molecules. With only a few exceptions, the improved potentials lie above the Morse curves and the corrections for moderately large internuclear separations may amount to ten percent of the energy of dissociation. Our treatment is based on the work of Dunham and the analysis of Coolidge, James and Vernon.

IN this paper the five spectroscopic constants which are available for most of the common diatomic molecules are used to improve the usual three-parameter Morse curves.[1] The two parameters in our correction term are easy to determine. In those cases where accurate potentials are known, our functions are satisfactory. With only a few exceptions, the improved potentials lie above the Morse curves and the corrections for moderately large internuclear separations may amount to ten percent of the energy of dissociation. The corrections are so large for some of the molecules that we are led to speculate on the existence of a hump or maximum in the potential energy. The higher vibrational energy levels are not known for most molecules, so that it is not possible to determine the potential uniquely for large internuclear separations from the existing experimental data. For these distances, it may be feasible to compute the potentials from the first- and second-order quantum-mechanical perturbation schemes. Our treatment is based on the work of Dunham[2] and the analysis of Coolidge, James and Vernon.[3]

The vibrational and rotational energy levels of a diatomic molecule are usually expressed in the form of a double infinite series

$$E_{v,K}=\sum_{i,j}Y_{ij}(v+\tfrac{1}{2})^i K^j(K+1)^j, \qquad (1)$$

where v is the vibrational and K is the rotational quantum number. The Y_{ij} are the usual spectroscopic constants and are better known in the nomenclature:

$$\begin{aligned} &Y_{00}=D, \quad Y_{10}=\omega_e, \quad Y_{20}=-\omega_e x_e, \\ &\qquad\qquad Y_{30}=\omega_e y_e, \quad Y_{40}=-\omega_e z_e, \\ &Y_{01}=B_e, \quad Y_{11}=-\alpha_e \quad \cdots, \\ &Y_{02}=-D_e, \quad Y_{12}=-\beta_e \quad \cdots. \end{aligned} \qquad (2)$$

Our problem is to determine the potential energy, $V(r)$, as a function of the internuclear separation r from a knowledge of these experimental constants. To do this we must assume a mathematical form for $V(r)$ which contains a number of parameters. These parameters are then adjusted to give the best fit to the observed spectroscopic constants.

The Morse curve has been used quite generally because of the simplicity of its form and the convenience of obtaining the energy levels in closed form. It may be written:

$$V_M=D[1-\exp(-x)]^2, \qquad (3)$$

where

$$x=2\beta\xi; \quad \xi=(r-r_e)/r_e;$$

β is a parameter for the Morse curve which has

* Presented at the Symposium on the Structure of Molecules and Aggregates of Molecules at the Fifth Annual Symposium of the Division of Physical and Inorganic Chemistry of the American Chemical Society, Columbia University, New York, December 30, 1940 to January 1, 1941.

[1] P. M. Morse, Phys. Rev. **34**, 57 (1929).
[2] J. L. Dunham, Phys. Rev. **41**, 713, 721 (1932).
[3] A. S. Coolidge, H. M. James and E. L. Vernon, Phys. Rev. **54**, 726 (1938).

the value

$$\beta=\omega_e/4(B_eD)^{\frac{1}{2}}$$

and r_e is the equilibrium separation. When the rotational quantum number is zero, the Schrödinger equation for this potential can be integrated exactly to give

$$\omega_e x_e=\omega_e^2/4D, \quad \omega_e y_e=\omega_e z_e=\cdots=0. \tag{4}$$

Pekeris[4] has calculated the rotational energy by a perturbation method. He found that to a good approximation:

$$\alpha_e=6B_e x_e\left[\left(\frac{B_e}{x_e\omega_e}\right)^{\frac{1}{2}}-\frac{B_e}{x_e\omega_e}\right], \quad D_e=\frac{4B_e^3}{\omega_e^2}. \tag{5}$$

Often spectroscopists assume *a priori* that $D_e=4B_e^3/\omega_e^2$ and fit their curves accordingly. Actually this form for D_e should only be true if the coupling between the rotation of the molecule and the spins and motions of the electrons is negligible. Deviations from this relation should be of interest but most of the published spectroscopic constants are computed in such a manner as not to show them. When the special relations of Eqs. (4) and (5) are not satisfied, it is obvious the Morse curve must be corrected.

I. Dunham's Method

The potential energy of a diatomic molecule may be expressed in the form:

$$V=a_0\xi^2[1+a_1\xi+a_2\xi^2+a_3\xi^3+\cdots]hc, \tag{6}$$

where $\xi=(r-r_e)/r_e$, h is Planck's constant and c is the velocity of light. This expansion converges rapidly for small values of ξ and is therefore useful in calculating the energy of the lowest vibrational states. This potential should be used in the Schrödinger equation to determine the vibrational and rotational energy levels in terms of the constants a_n. But the difficulties inherent in the integration of the Schrödinger equation for general values of the a_n made it necessary for Dunham[2] to use the WBK method. To this approximation he found that:

$$a_0=\omega_e^2/4B_e, \tag{7}$$

$$a_1=-1-\alpha_e\omega_e/6B_e^2, \tag{8}$$

$$a_2=\frac{5}{4}a_1^2-\frac{2}{3}\frac{x_e\omega_e}{B_e}, \tag{9}$$

$$a_3=\frac{\gamma_e\omega_e^2}{30B_e^3}-2a_1+\frac{3}{5}a_2+\frac{13}{5}a_1a_2-\frac{3}{2}(a_1^2+a_1^3)-1, \tag{10}$$

$$a_4=\frac{(y_e\omega_e)\omega_e}{5B_e^2}+\frac{7}{2}a_1a_3+\frac{17}{20}a_2^2-\frac{45}{8}a_1^2a_2+\frac{141}{64}a_1^4. \tag{11}$$

Two other equations show that $\omega_e c$ is the classical fundamental frequency for the potential $V=a_0hc\xi^2$ and that $B_e=h/(8\pi^2\mu r_e^2 c)$ where μ is the reduced mass of the molecule. One curious feature of these equations is that the coupling between the vibrational and rotational states manifests itself in a_1; while the first anharmonic correction in the vibrational energy levels, $\omega_e x_e$, does not appear until a_2. This has some interesting consequences in the selection of an appropriate potential function, as we shall show later. Dunham[2] subsequently improved upon the WBK approximation by adding corrections good to the third power of h. However, the numerical values of the correction terms are so small that we may neglect them in this treatment. For example, in the case of the ground state of hydrogen where the correction has its maximum effect, we find that ω_e should differ from its classical value by only one part in 6000 and B_e differs by one part in 12,000 from the above definition.

To test further the accuracy of WBK approximation, one may expand the Morse function to obtain the a_n's; then, substitute these a_n's into Dunham's expressions for the spectroscopic constants. The vibrational constants ω_e, $\omega_e x_e$, $\omega_e y_e$, and $\omega_e z_e$ obtained in this manner agree exactly with the expressions of Eq. (4) obtained by actual integration of the Schrödinger equation. While α_e calculated in this manner, agrees with the correct expression (5), the other rotational constants show some disagreement. This may be caused either by the approximations in the usual integration of the Schrödinger

[4] C. L. Pekeris, Phys. Rev. **45**, 98 (1934).

equation as discussed by Pekeris[4] or it may be caused in part by the WBK approximation. In any case we conclude that the WBK method is satisfactory for use in the determination of accurate potential energy functions. Experimental errors in the constants are larger in most cases than any errors which could be introduced by the use of this method.

II. The Five-Constant Potential Functions

Any acceptable form for a potential energy function should have a large value (if not infinite) as the nuclei come together; it should pass through a minimum at the equilibrium separation; and it should approach the energy of dissociation as the nuclei become far apart. At large separations, the slope of the function should be in agreement with the van der Waals forces but this is a refinement which has not received attention up to this time. There are an infinite number of functions which have a shape suitable for potential functions. Of these, the Morse function is the simplest and most generally useful. The three-parameter Manning and Rosen[5] function,

$$V=-\frac{A\exp(-cr)+B\exp(-2cr)}{[1-\exp(-cr)]^2} \tag{12}$$

has all of the general attributes of a good potential function and even gives the vibrational energy levels in closed form. But this function has never been generally adopted because most people prefer to use functions with greater flexibility, i.e., more parameters. The Pöschl-Teller[6] function has the same vibrational levels as the Morse function but a fourth parameter makes it possible to give a better adjustment to the rotational levels. It has the form:

$$V=D\left[\frac{\sinh^4 a}{\sinh^2(a+\beta\xi)}-\frac{\cosh^4 a}{\cosh^2(a+\beta\xi)}\right]. \tag{13}$$

The potential curve of Hylleraas[7] gives excellent agreement with the experimental energy levels but it has two disadvantages. It has six parameters, which is one more than we usually have experimental data to fit, and it is difficult to determine the constants. The Hylleraas potential has the form:

$$V=D\left[1-\frac{[1+a][1+b][\exp(2\beta'\xi)+c]}{[\exp(2\beta'\xi)+a][\exp(2\beta'\xi)+b][1+c]}\right]^2,$$

where a, b, c, β', D and of course, the r_e in ξ are the adjustable constants. The extended Morse curve of Coolidge, James and Vernon[3] appears more useful. It has the form:

$$\frac{V}{D}=\sum_{n=2,3,\cdots} C_n[1-e^{-2\beta'\xi}]^n. \tag{14}$$

In actual computations, Coolidge, James and Vernon have taken 7 terms in this series. This form of potential is flexible but it has one disadvantage, β' is difficult to determine and in fact loses its significance as we take more terms in the extended Morse curve. The three-term extended Morse function which uses five parameters is quite comparable in accuracy with the other five-parameter functions which we consider. The greater difficulty in evaluating these parameters makes it slightly less desirable.

For most diatomic molecules the four spectroscopic constants ω_e, $\omega_e x_e$, B_e, and α_e are known and listed in convenient tables. In addition the heats of dissociation are known from thermochemical or spectroscopic data. Therefore, we sought a potential function with five parameters which could be determined easily from the spectroscopic constants. Three of these constants could be used to determine the usual Morse function leaving two parameters to describe the corrections to the Morse curve. We find the following function satisfactory:

$$V=D[(1-e^{-x})^2+cx^3e^{-2x}(1+bx)]. \tag{15}$$

Here

$$x=2\beta\xi=\frac{\omega_e}{2(B_eD)^{\frac{1}{2}}}\left[\frac{r-r_e}{r_e}\right]$$

as in Eq. (3). The constants b and c are determined by the relation:

$$c=1+a_1(D/a_0)^{\frac{1}{2}}, \tag{16}$$

$$b=2+\left[\frac{7}{12}-\frac{Da_2}{a_0}\right]\Big/c, \tag{17}$$

[5] M. F. Manning and N. Rosen, Phys. Rev. **44**, 953 (1933).
[6] G. Pöschl and E. Teller, Zeits. f. Physik **83**, 143 (1933).
[7] E. A. Hylleraas, Zeits. f. Physik **96**, 661 (1935).

TABLE I.

	ω_e (CM^{-1})	$x_e\omega_e$ (CM^{-1})	B_e (CM^{-1})	α_e (CM^{-1})	r_e (A)	ENERGY OF DISSOC. (KCAL.)	D (CM^{-1})	a_0 (CM^{-1})	$-a_1$	a_2	2β	c	b
Hydrogen													
H_2	4405.3	125.325	60.872	3.0671	0.7414	103.23	38,292	79,703	1.60774	1.85848	1.4426	−0.1145	−0.704
HD	3817.1	94.958	45.668	1.9931	0.7413	104.06	38,292	79,762	−1.60798	1.84578	1.4433	−0.1141	−0.654
D_2	3118.8	64.15	30.429	1.0492	0.7417	105.03	38,292	79,915	1.58901	1.75071	1.4446	−0.1000	−0.556
$H_2(^3\Sigma_g^+)$	2665.9252	72.79390	34.22051	1.694884	0.9888	66.09	24,460	51,913	1.64090	1.95250	1.45672	−0.059533	−0.6543
Metal hydrides													
ZnH	1607.6	55.135	6.6794	0.2500	1.5947	19.6	7660	96,729	2.50138	2.31816	3.5536	0.2961	0.6501
CdH	1430.7	46.3	5.437	0.218	1.762	15.63	6184	94,119	2.75847	3.83430	3.901	0.2929	0.869
HgH	1432.7	108.8	5.58	0.373	1.741	8.58	3695	91,964	3.86052	5.63068	4.844	0.3099	0.671
Hydrogen halides													
HF	4141.305	90.866	20.967	0.879	0.9166	147.9	53,826	204,493	2.38007	4.19177	1.949	−0.2212	−0.351
HCl	2988.95	51.65	10.5909	0.3019	1.2747	102.16	37,244	210,884	2.34080	3.59797	2.379	0.0161	5.236
HBr	2649.67	45.21	8.471	0.226	1.414	86.73	31,674	207,200	2.39085	3.58717	2.558	0.0653	1.466
HI	2309.53	39.73	6.551	0.183	1.604	70.38	24,628	203,554	2.64138	4.67794	2.8098	0.0599	2.153
Halogens													
Cl_2	564.9	4.0	0.2438	0.0017	1.989	57.2	20,298	327,227	3.69278	6.10788	4.015	0.0803	−0.545
ICl	384.18	1.465	0.11414	0.000502	2.321	49.64	17,563	323,274	3.46724	6.47045	4.291	0.1920	0.793
Br_2	323.2	1.07	0.08091	0.00027	2.284	45.44	16,063	322,760	3.22156	4.15668	4.482	0.2812	0.661
I_2	214.36	0.593	0.03736	0.00012	2.667	35.56	12,550	307,483	4.07157	10.14034	4.950	0.1775	1.046
Oxides													
O_2	1580.36	12.073	1.4456	0.0158	1.2076	117.2	41,802	431,921	2.99143	5.61814	3.2145	0.0694	1.429
CO	2168.2	13.04	1.9310	0.01744	1.1284	210.8	74,840	608,634	2.69017	4.54428	2.8518	0.0567	1.568
NO	1906.52	14.504	1.709	0.0183	1.150	122.0	43,644	531,717	2.99094	5.52423	3.490	0.1430	1.092
SO	1123.73	6.116	0.70894	0.00562	1.4935	92.27	32,842	445,302	3.09425	6.21665	3.6823	0.1597	1.219
OH	3727.95	78.15	18.862	0.693	0.9710	99.0	36,548	184,201	2.21025	3.34436	2.2450	0.0155	7.181
Alkalies													
Li_2	351.346	2.557	0.67293	0.00719	2.6723	26.3	9379	45,861	1.92976	2.12179	2.211	0.1272	0.825
Na_2	159.23	0.726	0.15471	0.00079	3.079	18.0	6379	40,971	1.87592	1.27041	2.534	0.2597	0.516
K_2	92.64	0.354	0.05622	0.000219	3.923	11.85	4193	38,163	2.06982	1.15739	3.017	0.3139	0.547
Miscellaneous													
N_2	2359.61	14.445	2.007	0.018	1.095	170.2	60,738	693,542	2.75738	4.70575	3.379	0.1840	1.070
C_2	1641.55	11.67	1.6334	0.0149	1.3121	83	29,865	412,436	2.52793	3.22498	3.718	0.2725	1.017

where a_0, a_1 and a_2 are the Dunham coefficients of Eq. (6). Table I gives the values of b and c for 25 common diatomic molecules.

In making these calculations we were limited to those molecules for which the energy of dissociation is known reasonably accurately. The spectroscopic constants are those given by Herzberg.[8] The energies of dissociation are taken from Bichowsky and Rossini[9] after subtracting the rotational, vibrational, and translational heat contents at 18°C to get the energy of the molecules at absolute zero. The energies of dissociation of C_2 and of CO are still uncertain because of the heat of sublimation of graphite. We use the values recommended by Herzberg.

In all of those cases considered except Cl_2 the potential becomes large and positive when the nuclei come together. This spurious minimum for Cl_2 is completely unimportant since it does not appear for any physically attainable nuclear separation. For example, this potential gives an energy equal to the energy of dissociation when the nuclei are only 0.06A closer than the corresponding separation given by the Morse curve. However, all argument could be avoided in this case by using the potential:

$$V=D[(1-\exp(-x))^2 + cx^3(1+x)\exp(-(3-b)x)]. \quad (18)$$

This function does not have the spurious minimum and still it has all of the other desirable properties of Eq. (15). For some molecules the potential of Eq. (18) seems to be a little more accurate. But since this function is not generally applicable and in the other cases the difference between (15) and (18) is always small, we shall not use (18) for any general considerations.

Our confidence in the potential of Eq. (15) is bolstered by comparing it with the few accurate potentials for diatomic molecules which are available—the $1s\sigma 2s\sigma^3\Sigma_g^+$, state of H_2, the ground state of CdH, and the ground state of N_2.

[8] G. Herzberg, *Molecular Spectra and Molecular Structure, I* (Prentice-Hall, 1939).

[9] F. R. Bichowsky and F. D. Rossini, *Thermochemistry of Chemical Substances* (Reinhold Publishing Co., 1936).

1. The $1s\sigma 2s\sigma\,^3\Sigma^+_g$ state of H_2

Coolidge, James and Vernon[2] made use of the extensive experimental data available for this state to examine fourteen different types of potential functions and to compare them with an excellent nine-constant Morse function which fits the energy levels almost perfectly. Our five-parameter function (15) is better than any of the functions which they considered except a six-parameter Hylleraas function and a specially fitted Pöschl-Teller function. At $r=1.11$A ($x=0.1811$) our function (15) is only 0.017 kcal. higher than their nine-parameter function and at $r=1.43$A ($x=0.6491$) it is only 0.15 kcal. higher. At the same separation the usual Morse function is 0.043 and 0.34 kcal. higher, respectively. A five-parameter extended Morse function,[10] i.e., three terms of Eq. (14), gives an energy which agrees almost perfectly with the nine-parameter extended Morse function at $r=1.11$A, but is 0.11 kcal. lower than the nine-parameter function at $r=1.43$A. The three-term extended Morse function is thus about as far below the correct potential as (15) is above it.

2. The CdH molecule

All of the vibrational levels are known for CdH and Rydberg[11] constructed an accurate potential function to fit the experimental data. Rydberg[11] used a graphical method based on the WBK approximation. Subsequently Hylleraas[12] fitted a six-parameter potential curve by Dunham's method and obtained excellent agreement with Rydberg's potential. In Fig. 1 we have plotted the Rydberg-Hylleraas potential, the three-term extended Morse curve of the Coolidge, James and Vernon type (C.J.V.) and our potential (15). The constants for the extended Morse curve are

$$2\beta'=5.1067,\ C_2=0.5836,\ C_3=0.2684,\ C_4=0.1480.$$

The maximum deviation of our potential from the Rydberg curve is only 0.35 kcal. whereas the Morse function deviates by as much as 1.8 kcal. The extended Morse curve gives about as good agreement as our potential, the principal difference being that our correction is slightly too large while theirs is slightly too small (falls 0.30 kcal. below the Rydberg-Hylleraas curve). Actually our potential (18) agrees with the Rydberg-Hylleraas potential to within the width of the ink lines. Our only reason for not stressing (18) is that for some molecules it gives extremely large corrections which do not seem physically reasonable. All of these five-parameter functions therefore agree and are satisfactory for CdH.

3. The normal N_2 molecule

Hylleraas[10] has fitted the first 22 vibrational levels of N_2 with his six-parameter function. In Fig. 2 we have plotted his potential. The maximum deviation of the Morse curve from the Hylleraas curve is 14 kcal. For this case potentials (15) and (18) are almost identical. Both agree almost perfectly with the Hylleraas function although again the deviation from the Morse curve is large. The five-parameter extended Morse curve (C.J.V.) again falls below the Hylleraas curve, being 4.39 kcal. too low at $r=1.743$A. ($x=2$). The constants for this curve are

$$2\beta'=4.1939,\ C_2=0.6492,\ C_3=0.2223,\ C_4=0.1285.$$

It appears that the extended Morse curve (14) and our potential (15) are of about equal accuracy, but the constants are more readily determined for Eq. (15). Thus we gain confidence that our corrections are real rather than illusory.

In Table II we have used Eq. (15) to calculate the magnitude of the corrections to the Morse

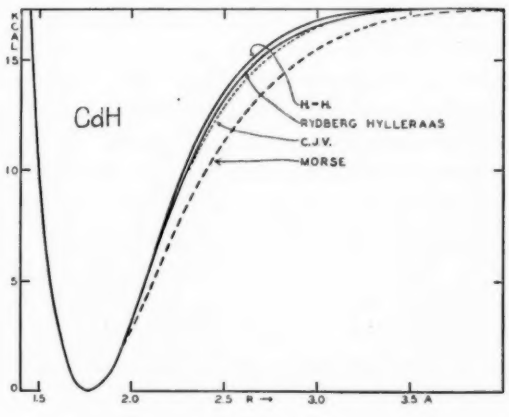


FIG. 1.

[10] We have taken $2\beta'=1.4858$; $C_2=0.9614$; $C_3=-0.1004$; $C_4=0.1390$ in Eq. (14).

[11] R. Rydberg, Zeits. f. Physik **73**, 376 (1932); **80**, 514 (1933).

[12] E. A. Hylleraas, Zeits. f. Physik **96**, 661 (1935).

curves for three different corresponding internuclear separations $x=1$, $x=2$, $x=3$ where the Morse function gives 40 percent, 75 percent, and 90 percent of the energy of dissociation, respectively. Figure 3 indicates where these points occur. Let us consider the molecules individually.

4. The hydrogen molecule

The potential curves for H_2, HD, and D_2 should be nearly the same. The differences, as

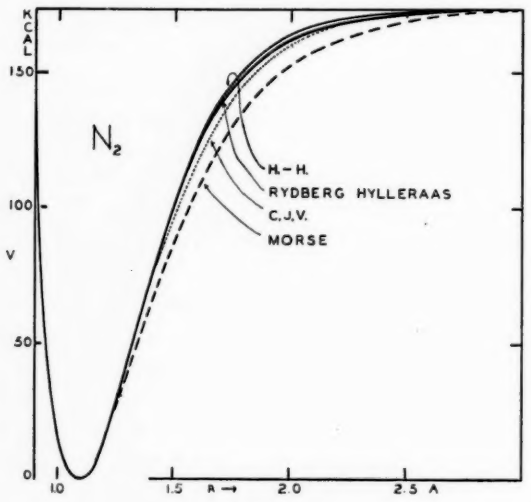


FIG. 2.

shown in Tables I and II, are caused primarily by the experimental errors which occur normally with the two-meter vacuum spectrograph used to photograph the spectra.[13] According to Professor Dieke,[14] more accurate experimental data should be forthcoming. These discrepancies show the order of accuracy of a_0, a_1, and a_2 which are probably present for other molecules in our table.

The corrections to the Morse curve for hydrogen are slight. The coupling between the rotation and vibration, α_e, indicates that the true potential should rise less steeply than the Morse curve; but the first anharmonic correction to the vibrational levels, $\omega_e x_e$, indicates that the Morse curve does not rise steeply enough. Fortunately these facts are not incompatible since α controls a_1 and $\omega_e x_e$ controls a_2. The correction term in Eq. (15) starts out negative at r_e and becomes positive for large separation of the nuclei. Qualitatively this must be the proper behavior but we cannot be sure that Eq. (15) handles this complicated effect quantitatively. Actually this function is an improvement on the Morse curve for small values of the internuclear separations and also for the large values but in the intermediate range the correction term may not change sign at the proper point. More work on this potential should be carried out as soon as the more accurate experimental data is available.

5. The metal hydrides

The potential curves for the metal hydrides rise much more steeply than the Morse function as is illustrated in Fig. 1 for CdH. Professor Mulliken[15] has warned us to look out for the possibility of a maximum and a long distance minimum in the potential curves for ZnH, CdH, and HgH due to the tendential approach of two interacting states. While Eq. (15) does not lead to these maxima, it is easy to make a slight change in the potential which does give the

TABLE II.

MOLECULE	V/D $x=1$	V/D $x=2$	V/D $x=3$	MAXIMUM CORRECTION KCAL.
Hydrogen				
H_2	−0.0046	0.0068	0.0085	<1
HD	− .0053	.0051	.0073	<1
D_2	− .0060	.0016	.0045	<1
Metal hydrides				
ZnH	.0661	.0998	.0585	2.2
CdH	.0741	.1175	.0707	2.1
HgH	.0701	.1063	.0625	1.1
Hydrogen halides				
HF	− .0194	−.0096	.0008	−3.0
HCl	.0136	.0271	.0180	2.9
HBr	.0218	.0376	.0236	3.4
HI	.0256	.0466	.0299	3.4
Halogens				
Cl_2	.0049	−.0011	−.0034	<1
ICl	.0466	.0727	.0434	3.6
Br_2	.0632	.0957	.0561	4.4
I_2	.0491	.0804	.0491	2.9
Oxides				
OH	.0170	.0346	.0231	3.6
CO	.0197	.0344	.0216	7.3
O_2	.0228	.0392	.0245	4.7
NO	.0405	.0667	.0409	8.3
SO	.0479	.0804	.0497	7.6
Alkalies				
Li_2	.0314	.0494	.0296	1.3
Na_2	.0533	.0773	.0443	1.4
K_2	.0657	.0963	.0555	1.2
Miscellaneous				
C_2	.0744	.1211	.0739	10.3
N_2	.0515	.0846	.0518	14.7

[13] C. R. Jeppeson, Phys. Rev. **49**, 797 (1936).
[14] G. H. Dieke, Phys. Rev. **47**, 661 (1935).
[15] Robert S. Mulliken, J. Phys. Chem. **41**, 5 (1937).

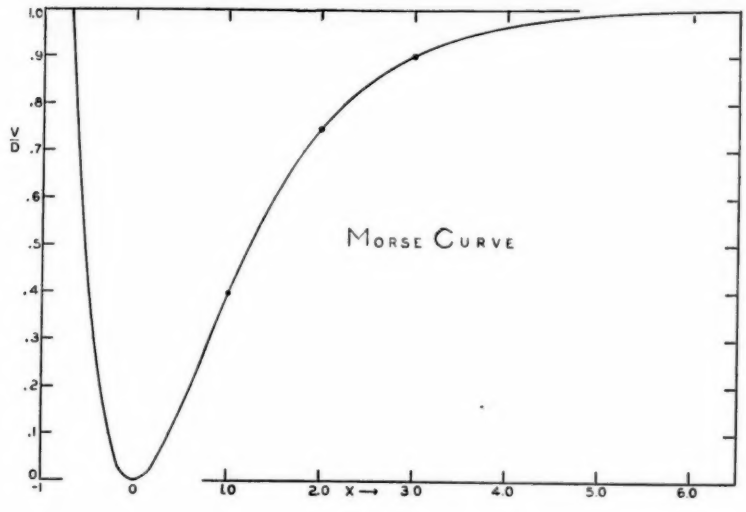


FIG. 3.

maximum and is still consistent with all of the experimental information.

7. The hydrogen halides

The corrections to the Morse curve for the hydrogen halides increase as we go down the series HF, HCl, HBr, and HI. However, only the lowest few vibrational levels are known for these molecules. The corrections to the Morse curve for HF are small and may not be significant. This may be another example showing that the Morse curve works best for those molecules which do not have inner shells.

The potential for HCl is shown in Fig. 4. Here the correction is fairly small and agrees with that which we would expect on the basis of the three-term extended Morse curve.

$2\beta' = 2.6089$, $C_2 = 0.8319$, $C_3 = 0.0855$, $C_4 = 0.0826$.

The potential curves for HBr and HI behave perfectly normally. The maximum corrections are 3.5 and 4.5 kcal., respectively. The large anharmonicity of the HI molecule accounts in part for the difficulty of computing the equilibrium constants for $H_2 + I_2 \rightleftarrows 2HI$ from the spectroscopic data.[16,17]

Since only the three lowest vibrational levels of HI are known, there may be considerable experimental error in ω_e, $\omega_e x_e$ and hence in our correction term.

8. The halogens

The potential curves for ICl, Br_2, and I_2 rise more steeply than the Morse curve, have maximum corrections of three or four kcal., but otherwise behave quite normally, as seen in Fig. 5. The potential hump for ICl which is indicated from the absorption spectra and predissociation work of Brown does not appear from our potential.[15]

The potential for Cl_2 deviates very little from the Morse function.

9. The alkalies

A great deal of experimental information is available for the alkalies and it would be desirable to construct very accurate potentials for them. According to Eq. (15), the true potentials rise more steeply than the Morse curves, have maximum corrections to the Morse curves of about a kilocalorie, and seem to behave perfectly normally. Our potential for Na_2 gives a fifth vibrational energy level which differs from the experimental value by only 0.015 kcal.

10. The oxides and OH

Here the corrections to the Morse curve are quite large but the potentials appear quiet normal.

[16] G. M. Murphy, J. Chem. Phys. **4**, 344 (1936).

[17] Our energy of dissociation of HI was obtained from some unpublished data of Crist which he gave us in a private communication.

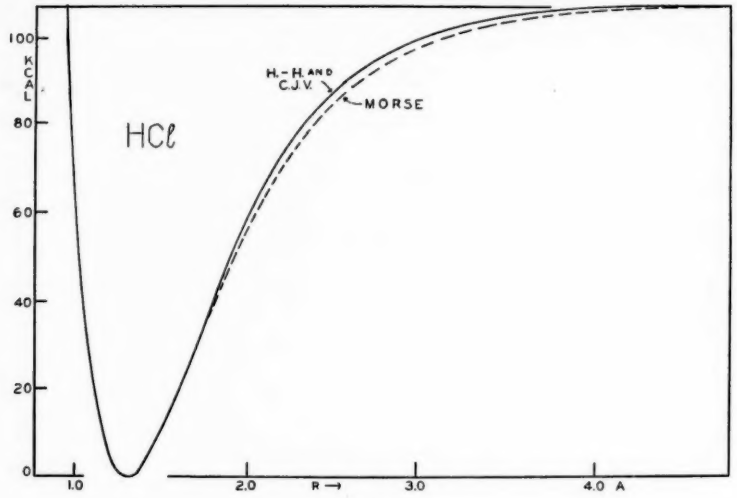


FIG. 4.

11. C_2 and N_2

The C_2 molecule appears normal with a rather large correction to the Morse curve. The molecule N_2 has been discussed previously. The corrections to the Morse curve are very large at intermediate distances, but Eq. (15) reproduces the accurate Hylleraas potential almost exactly.

IV. Significance of the Spectroscopic Constants

In this treatment we adjusted the potential function to fit the spectroscopic constants. These constants are readily available and our method is convenient. But as Coolidge, James and Vernon have pointed out,[3] greater accuracy can be obtained by making a least square fit of the individual energy levels. The reason for this is that the spectroscopic constants contain both random experimental errors and systematic errors resulting from the particular polynomial formula which the energy levels are supposed to obey. Often (as in the case of N_2) it is found that one polynomial is adequate for expressing the lower vibrational levels, whereas an altogether different polynomial is required for the higher states. This phenomena may be due to the fact that a true expression for the energy levels might contain an exponential, as for example:

$$E_v=A[1-\exp(-\tfrac{1}{2}(v+\tfrac{1}{2}))]\times[1+0.1(v+\tfrac{1}{2})-0.005(v+\tfrac{1}{2})^2].$$

For small values of $(v+\frac{1}{2})$ the energy levels would be given by the power series:

$$E_v=0.5(v+\tfrac{1}{2})-0.075(v+\tfrac{1}{2})^2+\cdots$$

but for large quantum numbers, the power series is:

$$E_v=1+0.1(v+\tfrac{1}{2})-0.005(v+\tfrac{1}{2})^2+\cdots.$$

The reason for this difference is that exponentials of large negative numbers are asymptotically zero. Thus it is never safe to trust the accuracy of power series developments.

This same criticism can be made of the Dunham method for obtaining the potential functions. It is obvious that the power series expansion of the potential does not converge

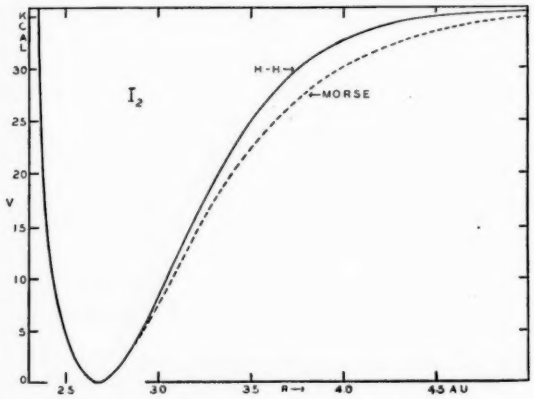


FIG. 5.

rapidly for large internuclear distances and therefore Dunham's method is not adapted to the calculations of the higher vibrational levels. For this purpose we would suggest either one of two methods. First, one could substitute the assumed potential directly into the Schrödinger equation and carry out the integrations numerically to find the corresponding energy levels. Or, one could develop a perturbation method using the Morse potential and the Morse wave functions for the unperturbed state. The integrals of Pekeris[4] would make this method easy to carry out. At the present time these refinements are not warranted by the accuracy of the spectroscopic data for most molecules.

Thus we have made a rough survey of the corrections which must be made to the Morse curve for all of the common diatomic molecules. In some cases the corrections are large. As Mulliken[15] has pointed out, there does not seem to be any universal form for a potential energy function, each molecule having its own peculiar properties. Significant advances in the determination of molecular potential functions will come through closer coöperation of the quantum mechanician and the spectroscopist.

One of us (J.O.H.) would like to express his gratitude to the Wisconsin Alumni Research Foundation for financial aid given to him at the beginning of this work.

JANUARY, 1941 JOURNAL OF CHEMICAL PHYSICS VOLUME 9

Force Constants in Some Organic Molecules*

BRYCE L. CRAWFORD, JR.** AND STUART R. BRINKLEY, JR.
Sterling Chemistry Laboratory, Yale University, New Haven, Connecticut

(Received October 9, 1940)

Consistent normal coordinate treatments, involving force-constants which are related to bond structures and which may be transferred from one molecule to another, are applied to hydrogen cyanide, methyl cyanide, and the methyl halides. The connection with previous treatments of acetylene, ethane, and methyl and dimethyl acetylene is discussed. A method of setting up such a consistent treatment is described, and a table of force constants for a number of bond structures is given. The structural significance of these force constants is briefly discussed.

IN analyzing the vibrational spectra of polyatomic molecules, it has become increasingly common to carry out normal coordinate treatments. One important reason for such a treatment is the aid thus given in fixing the fundamental frequencies of complicated molecules. When a few fundamentals are unambiguously assigned, as through symmetry selection rules or from polarization data, it often becomes possible to locate less certain frequencies, at least approximately, by means of a normal coordinate calculation in which the force constants are evaluated from the unambiguous assignments. In view of the difficulty of vibrational analyses for complex molecules, the importance of such aids can scarcely be overemphasized.

Of more fundamental importance are the magnitudes of the force constants which are evaluated. If a valence-force potential is used, these force constants are an additional property of the bonds involved in the molecule, and they may be used to study the character of different bonds and their variation in different molecules. The connection which has been found between the force constant of a diatomic molecule and its bond length[1] shows us the type of correlation we may expect.

---

* Presented at the Symposium on the Structure of Molecules and Aggregates of Molecules at the Fifth Annual Symposium of the Division of Physical and Inorganic Chemistry of the American Chemical Society, Columbia University, New York, December 30, 1940 to January 1, 1941.

** Present address: School of Chemistry, University of Minnesota, Minneapolis, Minnesota.

[1] R. M. Badger, J. Chem. Phys. **2**, 128 (1934); C. H. Douglas-Clark, Phil. Mag. **18**, 459 (1934); and later papers.

If we consider a series of molecules involving the same bonds, we may expect the bond force constants to remain unchanged (or at worst to change very little) from one molecule to another. It is essential that the same type of potential function be used for each molecule, and that it be sufficiently general to include all important force constants.[2] Eventually, we may hope to have a table of bond force constants, similar to tables of bond energies, such that if we know the bond structure of a given molecule, we may calculate the complete set of its fundamental frequencies *a priori*, with an accuracy sufficient for such purposes as thermodynamic calculations. This of course represents a quantitative formulation of the idea of bond frequencies, which has long been helpful in vibrational analysis.

In the present paper we are investigating this possibility and seeking to find to what extent force constants are transferable from molecule to molecule. We find that in the 6 related molecules considered, hydrogen cyanide, acetylene, ethane, methyl acetylene, dimethyl acetylene, and methyl cyanide, the different bond structures call for 20 distinct force constants; with these, the 52 fundamental frequencies are fitted satisfactorily. A separate series, the methyl halides, is also considered in order to test the transfer of force constants for less closely related structures.

Mathematical Technique

In setting up the secular equations for the various molecules, we have used the technique recently developed by Wilson.[3] We found this procedure quite straightforward and easily handled, and it is particularly convenient in the present case because certain sections of the matrices involved were identical for a series of molecules. The most laborious part of the calculation is the evaluation of the necessary minors. We found it most convenient to evaluate the smallest (two-rowed) minors first, then to use these in evaluating the three-rowed minors, and thus to proceed through the larger determinants using at each step the results already obtained.

The approximate factoring of the hydrogen stretching frequencies (3000 cm^{-1}), obtained by setting the appropriate force constants equal to infinity,[3] is extremely useful and quite accurate. When the other frequencies in a factor of the secular equation were under 2000 cm^{-1}, we found this approximation to be accurate to within 2 cm^{-1}. This technique is thus more accurate as well as more convenient than that previously proposed,[4] the increased accuracy being due to the fact that the perturbation correction is essentially included in Wilson's technique.

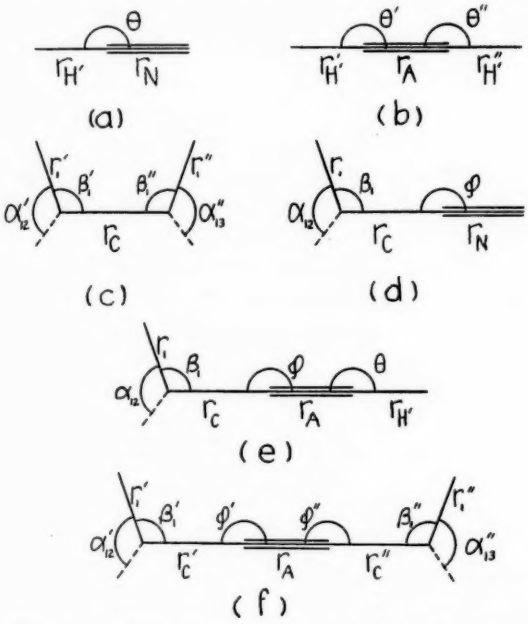


Fig. 1. Valence-force coordinates for the six related molecules discussed: (a) HCN, (b) C_2H_2, (c) C_2H_6, (d) CH_3CN, (e) methyl acetylene, (f) dimethyl acetylene. Dotted lines indicate projections of certain CH bonds onto the plane of the paper. Equilibrium values for all α's and β's are 109°28′, the tetrahedral angle; for all φ's and θ's, 180°.

Coordinates

The *valence-force coordinates* (bond lengths and angles) used are illustrated in Fig. 1. They were chosen for convenience in setting up the F matrix.[3] Symmetry factoring was obtained by using *valence-force symmetry coordinates* (VFSC)

[2] G. Glockler and F. T. Wall [J. Chem. Phys. **5**, 813 (1937)] have treated series of molecules, holding the force constants invariant. The treatment is not complete, however, nor are the potential functions as general as ours.

[3] E. B. Wilson, Jr., J. Chem. Phys. **7**, 1047 (1939); and Professor Wilson's paper in this Symposium. Reference should be made to these papers for the details of this technique.

[4] B. L. Crawford, Jr. and J. T. Edsall, J. Chem. Phys. **7**, 223 (1939).

which are orthogonal combinations of the valence-force coordinates.

In order to obtain bending and interaction force constants which are dimensionally uniform with the diagonal stretching constants, we chose the angle coordinates to be, not the angles themselves, but the following quantities: $r_H(\Delta\theta)$, $r_C(\Delta\varphi)$, $r_0(\Delta\alpha)$, and $r_0(\Delta\beta)$, where r_0 is the equilibrium CH distance in the CH_3 group. This choice affects the values of the force constants, of course. The transformations to other types of force constant can be made easily; the type used here is convenient to use, dimensionally uniform, and remarkably transferable between molecules, as is discussed below.

It will be noted that the CH_3 coordinates form a redundant set, since *six* angles (3 α's and 3 β's) are used. There is, then, a relation between these six angles, and we believe that the best way of eliminating the redundancy from the final secular equation is to take this relation to be itself a VFSC. Thus in the A_1 factor of CH_3CN, the redundant VFSC is

$$S_1=(6)^{-\frac{1}{2}}r_0(\sum_3\Delta\alpha_{ij}+\sum_3\Delta\beta_i)=0.$$

The other VFSC are then taken to be orthogonal to the redundant one. Since the redundant VFSC vanishes by definition, it will contribute a row and column of zeros to both the F and G matrices.[3] This method of eliminating the redundancy is easily carried out, avoids subsequent matrix manipulation or transformation of force constants, and yields F and G matrices (in terms of the VFSC) which are in large part transferable from one molecule to another.

Potential function

Having chosen our valence-force coordinates, and having thereby defined our G matrix, the definition of the F matrix remains. The most general potential function would be obtained by placing an arbitrary force constant in each element of the F matrix (subject to symmetry restrictions); this, however, involves too many constants to be practicable. It is necessary, therefore, to decide which interaction terms are to be included. We found it convenient to define our F matrices in terms of the original valence-force coordinates, R_i. Since these are bond lengths and angles, there will be a "bond-frequency" associated with each R_i, the magnitude of which is roughly known. We then used the following rule: *Setting up the F matrix in terms of valence-force coordinates R_i, set all force constants $F_{ij}=0$ except (1) diagonal force constants F_{ii}; (2) F_{ij} for which R_i and R_j have bond frequencies not too widely separated and for which either (a) $G_{ij}\neq 0$, or (b) if $G_{ij}=0$, some bond length (not angle) R_k exists such that $G_{ik}\neq 0$, $G_{jk}\neq 0$.* We believe that this rule gives a potential function in which all important interaction constants are included; it must, of course, be tested more thoroughly.

We use the letter K for a diagonal stretching force constant; H for a diagonal bending constant. Small letters are used for interaction

TABLE I. *Force constants for the valence-force coordinates. Values are given in units of 10^5 dynes/cm. The valence-force coordinates involved, except for the methyl halides, are shown in Fig. 1. The additional constants necessary in ethane,[5] to care for the interactions between the two methyl groups, are omitted here. The constant K_C was determined as a function of r_C; see text.*

CONST.	VALUE	COORDINATES INVOLVED
K_H	4.790	(Δr_i) in CH_3
H_α	0.46	$r_0(\Delta\alpha)$ in CH_3
$l_{\alpha\beta}$	+0.01	$r_0(\Delta\alpha)\cdot r_0(\Delta\beta)$ in CH_3
H_β	0.55	$r_0(\Delta\beta)$ in $C-CH_3$
$k_{C\gamma}$	−0.24	$(\Delta r_C)\cdot r_0(\Delta\alpha)$; $(\Delta r_C)\cdot r_0(\Delta\beta)$; see text
K_C		(Δr_C); see text
H_β	0.76	$r_0(\Delta\beta)$ in $F-CH_3$
$k_{F\gamma}$	−0.67	similar to $k_{C\gamma}$
K_F	5.96	C−F bond
H_β	0.58	$r_0(\Delta\beta)$ in $Cl-CH_3$
$k_{Cl\gamma}$	−0.49	similar to $k_{C\gamma}$
K_{Cl}	3.64	C−Cl bond
H_β	0.52	$r_0(\Delta\beta)$ in $Br-CH_3$
$k_{Br\gamma}$	−0.476	similar to $k_{C\gamma}$
K_{Br}	3.129	C−Br bond
H_β	0.45	$r_0(\Delta\beta)$ in $I-CH_3$
$k_{I\gamma}$	−0.456	similar to $k_{C\gamma}$
K_I	2.648	C−I bond
$K_{H'}$	5.850	$(\Delta r_{H'})$
K_A	15.587	(Δr_A)
K_N	17.727	(Δr_N)
H_φ	0.155	$r_C(\Delta\varphi)$
H_θ	0.2096	$r_{H'}(\Delta\theta)$
k_{tC}	+0.22	$(\Delta r_C)(\Delta r_A)$; $(\Delta r_C)(\Delta r_N)$
$k_{tH'}$	+0.028	$(\Delta r_{H'})(\Delta r_A)$; $(\Delta r_{H'})(\Delta r_N)$
$h_{\theta\theta}$	+0.0784	$r'_{H'}(\Delta\theta')r''_{H'}(\Delta\theta'')$
$h_{\varphi\varphi}$	+0.0670	$r'_C(\Delta\varphi')r''_C(\Delta\varphi'')$
$h_{\varphi\theta}$	+0.0725	$r_C(\Delta\varphi)r_{H'}(\Delta\theta)$
k_{CC}	+0.193	$(\Delta r'_C)(\Delta r_C'')$

TABLE II. *Ethane*, C_2H_6. C—H, *1.093;* C—C, *1.55A. Tetrahedral angles.*

	OBSERVED	CALCULATED[5]	PERCENT ERROR
A_{1g}	993	1010	+1.7
	1375	1375	—
	2925	2900	−0.9
A_{2u}	1380	1390	+0.7
	2925	2895	−1.0
E_u	827	830	+0.4
	1465	1495	+2.0
	2980	3010	+1.0
E_g	1170	1200	+2.6
	1460	1445	−1.0
	2960	3020	+2.0

constants: k for the interaction between two bond lengths (stretch-stretch) or between a bond length and an angle, h and l for interactions between two angles.

To illustrate the application of the proposed rule and the meaning of our symbols, we may write the potential function for the hydrogen cyanide molecule in the form

$$2V=K_{H'}(\Delta r_{H'})^2+K_N(\Delta r_N)^2 +H_\theta(r_{H'}\Delta\theta)^2+2k_{tH'}(\Delta r_{H'})(\Delta r_N).$$

The corresponding F matrix is

$$F=[2V]=\begin{bmatrix} K_{H'} & k_{tH'} & 0 \\ k_{tH'} & K_N & 0 \\ 0 & 0 & H_\varphi \end{bmatrix}.$$

If one applies the rule proposed above to the CH_3 group coordinates, one would include four types of interaction constants: one between opposite angles, $h_{\alpha\beta}$, and three between adjacent angles of various types, $l_{\alpha\alpha}$, $l_{\beta\beta}$, and $l_{\alpha\beta}$. Actually we have adopted a further simplification and taken only $l_{\alpha\beta}$ to be nonvanishing, following Stitt's potential for ethane.[5] This seems to be satisfactory in the cases here considered.

A given interaction constant occurs in two places in the F matrix, since $F_{ij}=F_{ji}$. Therefore it is necessary to use a factor of two in comparing the values given here with those previously published.[5–7]

RESULTS OF CALCULATIONS

The force constants used are tabulated in Table I, and the frequencies calculated for the various molecules are compared with the observed values in the succeeding tables. It must be borne in mind that an accuracy better than 2 percent is not to be expected, since the anharmonic terms will affect the frequencies by about this amount; even larger anharmonic corrections may be present in the hydrogen stretching frequencies.

Many of the force constants are taken from previous work; the source of each is given in the discussion which follows.

Previous work

Acetylene.—The normal coordinate treatment was originally carried out by Colby.[8] The constants were recalculated[6] to conform to our type of bending force constants and to the bond lengths C≡C, 1.204; C—H, 1.057A. Strict adherence to the rule proposed above would give four constants for the stretching frequencies: K_A, $K_{H'}$, $k_{tH'}$, and $k_{H'H'}$. Following Colby, we set the last equal to zero, a somewhat arbitrary procedure justified only by the results. The two bending frequencies required the constants H_θ and $h_{\theta\theta}$. Thus five constants were evaluated from the five frequencies of this molecule; the agreement is, of course, exact, and the comparison table is omitted. This potential function also gives satisfactory results for C_2D_2 and C_2DH.[8,9]

TABLE III. *Dimethyl acetylene*, $C_2(CH_3)_2$. C≡C, *1.204;* C—C, *1.462;* CH_3 *group as in ethane.*

	OBSERVED	CALCULATED	PERCENT ERROR
A_1	725	722	−0.4
	1380	1402	+1.6
	2270	2273	+0.1
	2916	2904	−0.4
$\tilde{A}_2$	1126	1124	−0.2
	1380	1402	+1.6
	2976	2904	−2.4
E	213	214	+0.5
	1050	1036	−1.3
	1468	1476	+0.5
	2976	3019	+1.4
$\tilde{E}$	371	370	−0.3
	1029	1040	+1.1
	1448	1476	+1.9
	2966	3019	+1.8

Ethane.—The normal coordinate treatment of Stitt[5] supplied us with our constants for the

[5] F. Stitt, J. Chem. Phys. **7**, 297 (1939).
[6] B. L. Crawford, Jr., J. Chem. Phys. **8**, 555 (1939).
[7] B. L. Crawford, Jr., J. Chem. Phys. **8**, 526 (1940).
[8] W. F. Colby, Phys. Rev. **47**, 388 (1935).
[9] F. Stitt, J. Chem. Phys. **8**, 56 (1940).

CH_3 group. The potential function used was discussed above; the constants involved here are $K_H, K_C, H_\alpha, H_\beta, l_{\alpha\beta}$, and $k_{C\gamma}$. Stitt used additional constants to express the interaction between the two adjacent CH_3 groups, the form of the interaction depending on the choice of D_{3d} or D_{3h} equilibrium configuration. We have omitted these constants since they are not involved in the molecules discussed here. The constant $k_{C\gamma}$ is the difference $(k_{C\alpha}-k_{C\beta})$, which always enters the equations in this form. The constant K_C was found by Stitt to be 4.50 for the C−C bond distance in ethane, 1.55A.[5] The C−H distance 1.093A and tetrahedral angles in the CH_3 group have been assumed throughout the present work. Table II gives the frequency comparison for C_2H_6. The potential function is also satisfactory for C_2D_6.[5]

Dimethyl acetylene.—The actual transfer of force constants from one molecule to another was begun in this case.[6] In addition to the constants from ethane and acetylene, four new constants H_φ, $h_{\varphi\varphi}$, k_{CC}, and k_{tC} were evaluated, and a new value $K_C=5.183$ for the C−C bond distance 1.462A found. The frequency comparison is given in Table III.

Methyl acetylene.[7]—In this molecule, the bond structure is so closely related to those of dimethyl acetylene and acetylene that only one new force constant, $h_{\varphi\theta}$, is required. This constant was arbitrarily set equal to the arithmetic mean of $h_{\theta\theta}$ and $h_{\varphi\varphi}$; thus all constants were evaluated *a priori*. The frequency comparison in Table IV shows how remarkably successful the calculation was.

New results

Hydrogen cyanide.—In accordance with the rule for finding necessary force constants, we should have in this molecule three constants for the stretching frequencies, $K_{H'}$, K_N, and $k_{tH'}$, and one bending constant H_θ for the bending frequency. The molecule is quite similar in structure to acetylene; we therefore took three of the constants from that molecule, and evaluated only one new constant, K_N. The constants were also tested on DCN; frequency comparisons for both molecules are given in Table V. Frequencies are taken from tables[10] and Bartunek and Barker.[11]

TABLE IV. *Methyl acetylene,* $(CH_3)C_2H$. C≡C, *1.204;* C—C, *1.462;* C−H′, *1.057;* CH_3 *as in ethane.*

	Observed	Calculated	Percent Error
A_1	926	925	−0.1
	1380	1402	+1.6
	2151	2137	−0.7
	2910	2905	−0.2
	3430	3331	−2.9
E	333	332	−0.3
	642	656	+2.3
	1041	1040	−0.1
	1444	1475	+2.1
	2995	3021	+0.9

TABLE V. HCN *and* DCN. C−H′, *1.06;* C≡N, *1.15A.*

		Observed	Calculated	Percent Error
HCN,	A	2089	2087	−0.1
		3290	3341	+1.6
	E	712	700	−1.7
DCN,	A	2630	2614	−1.0
		1906	1921	+0.8
	E	569	558	−1.9

Methyl cyanide.—This molecule is closely related to methyl acetylene on the one hand and HCN on the other; thus all but one of the force constants required by the rule could be transferred from other molecules as indicated in Table I. The C−C bond length is 1.49A;[12] accordingly, the force constant K_C should lie between that in ethane (bond length 1.55A) and that in methyl acetylene (1.462A). We interpolated between these two in the following way. We set up a two-constant relation giving K_C as a function of the bond length r_C, using, as a reasonable guess, the form proposed by Badger[1] for diatomic molecules; the two constants were evaluated from the force constants for ethane and methyl acetylene. The relation found is

$$K_C=33.68/(r_C+0.4061)^3,$$

where K_C is in units of 10^5 dynes/cm and r_C in angstroms. It will be noted that the "correction constant" 0.4061 differs in sign from Badger's similar constants for diatomic molecules. (This relation does not hold, with the same constants, for the C≡C bond.) Substituting $r_C=1.49$, we found K_C to be 4.94. With this value, all force

[10] H. Sponer, *Molekülspektren* (Berlin, 1935), Vol. I; Landolt-Börnstein Tabellen, Erg. III, 1194.
[11] P. F. Bartunek and E. F. Barker, Phys. Rev., **48**, 516 (1935).
[12] L. Pauling, H. D. Springall and K. J. Palmer, J. Am. Chem. Soc. **61**, 927 (1939).

TABLE VI. *Methyl cyanide,* CH_3CN. C≡N, *1.16;* C−C, *1.49A;* CH_3 *as in ethane.*

	OBSERVED*	CALCULATED	PERCENT ERROR
A_1	917	902	−1.6
	1370	1396	+1.9
	2265	2251	−0.6
	2942	2905	+1.2
E	376	378	+0.5
	1040	1036	−0.4
	1415	1475	+4.2
	3077	3022	−1.8

* It should be noted that these are *liquid-phase* frequencies.

constants are fixed *a priori*, and the frequencies given in Table VI were obtained.

Unfortunately, the experimental data for this molecule all pertain to the liquid phase;[13] in view of the well-known shift between liquid and gas phase frequencies, the experimental values are not as reliable as those available for the other molecules. This makes the pleasing agreement shown for seven out of eight frequencies in Table VI less significant; on the other hand, the large error of 4.2 percent in the case of the perpendicular hydrogen bending (1475 cm^{-1}) may also be regarded less seriously than if the 1415 value were a gas phase frequency. Were the principle of force constant transfer more firmly established, Table VI might be viewed as an indication that the gas-to-liquid shift for methyl cyanide is large only for the 1475 cm^{-1} fundamental; as a test of this principle, the table shows, we feel, satisfactory agreement. Gas phase data will supply a more revealing test.

Methyl halides.—In order to investigate the variation of force constants with bond type, we made calculations for the series of methyl halides, CH_3X. We hope also to use the constants so evaluated in a study of the ethyl halides. Designating the HCX angle as β, and the C−X bond as r_X, we found that six force constants were required by the rule: K_H, H_α, $l_{\alpha\beta}$, H_β, K_X, and $k_{X\gamma}$. Of these the first two may legitimately be transferred from ethane (Table I); we also made the approximation of taking $l_{\alpha\beta}$ from this table. The remaining three constants were fitted to each of the methyl halides; the values found are given in Table I. The frequency comparisons are given in Tables VII to X.

[13] The data are summarized and a reasonable assignment given by J. W. Linnett, J. Chem. Phys. **8**, 91 (1940).

Methyl fluoride showed less satisfactory agreement than the others, indicating that $l_{\alpha\beta}$ should be evaluated anew for this molecule. It should be remarked that $l_{\alpha\beta}$ is more important in CH_3F, since the α and β frequencies (1200 and 1476 cm^{-1}) are less widely separated than in the other methyl halides.

The constants for the methyl halides (H_β, K_X, and $k_{X\gamma}$) will depend to some extent on the values chosen for the C−X distances in these

TABLE VII. *Methyl fluoride,* CH_3F C−F, *1.385A;* CH_3 *as in ethane.*

	OBSERVED	CALCULATED	PERCENT ERROR
E	1200	1185	−1.25
	1476	1498	+1.5
	2987	3023	+1.2
A_1	1048	1048	—
	1460	1479	+1.3
	2920	2904	−0.5

molecules. These distances, unfortunately, are rather uncertain at present.[14] We have rather arbitrarily taken the values shown in Tables VII–X. A different set of distances would have affected the values of the constants; i.e., H_β, K_X, and $k_{X\gamma}$ would have to be evaluated anew to fit the lowest E frequency and the two lower A_1 frequencies in each case. The effect on the other frequencies would be negligibly small, even in the case of methyl fluoride.

DISCUSSION

In considering the constants listed in Table I, it must be remembered that they represent merely a preliminary set. We have as yet examined only a few molecules; moreover, a final table of force constants should consist of "best values," each constant being fitted to a number of different molecules, rather than evaluated from a single case as has been done here. It must also be remembered that, in view of the magnitude of anharmonic corrections, the number of significant figures given for many of the force constants is too large. They are given as being the values we actually used.

With these limitations, however, we believe that Table I gives a set of fairly reliable con-

[14] See G. B. B. M. Sutherland, Trans. Faraday Soc. **34**, 325 (1938).

stants, based upon a treatment of sufficient generality and consistency so that they can be transferred from one molecule to another. The constants may also be regarded as having a real physical significance, being properties of certain bonds and angles which can be used to study these structures.

Stretching constants

It is apparent from the table that the stretching constants (K's) vary not only with bond type but also, as might be expected, with bond distance. In the case of the C−C bond, we found that a relation of the Badger[1] type (using different constants from those given by Badger for diatomic molecules, however) was satisfactory as long as the bond was basically a single bond. The data are not yet sufficient to test such a relation for other bonds (e.g., K_H and $K_{H'}$ for the C−H bond).

The constants for the C−X bonds in the methyl halides show a regular variation, as one would expect. We have not examined this variation closely, preferring to wait until consistent force constants for a greater number of bonds are available.

Bending constants

These (H's) are rather more interesting than the stretching constants. They seem to be dependent only on the types of the bonds forming the angle, and not on the bond lengths; thus H_β in the C−CH_3 structure is constant for C−C lengths from 1.46 to 1.55.

More remarkable is the fact that H_θ and H_φ can be transferred from structures involving the C≡C bond to those involving the C≡N bond. This seems to agree with the theory of directed valence;[15] in cases such as H_θ and H_φ, we can regard the restoring force as arising almost exclusively from the "directed" nature of the bonds, and very little from forces between the atoms at the ends of the angle. Hence these constants will depend not on the end atoms, but on the bonds themselves, and will be the same if the orbitals involved in the two bonds are the same.

The values of H_β for the methyl halide series may be taken as showing the variation resulting when the bond type does change; here, however, the end atoms are closer together, and the forces acting between them must undoubtedly be taken into account.

Interaction constants

The significance of these constants is less obvious than that of the K's and H's; we shall merely remark that they, like the bending constants, seem to be sensitive to bond type and not to bond lengths.

In conclusion, we wish to emphasize again the preliminary nature of our table of force constants; we hope that it may be revised and extended to give a consistent set of constants which will be of real aid both to vibrational analyses and to the study of bond structures.

It is a pleasure to thank Professor E. B. Wilson, Jr., for helpful discussion and for making available to us the latest developments of his technique for finding the secular equations, before their publication.

TABLE VIII. *Methyl chloride*, CH_3Cl C−Cl, *1.66A*; CH_3 *as in ethane.*

	Observed	Calculated	Percent Error
E	1020	1019	−0.1
	1460	1478	+1.2
	3047	3021	+0.9
A_1	732	732	—
	1355	1355	—
	2920	2902	−0.6

TABLE IX. *Methyl bromide*, CH_3Br C−Br, *1.91A*; CH_3 *as in ethane.*

	Observed	Calculated	Percent Error
E	957	952	−0.5
	1450	1474	+1.7
	3061	3021	−1.3
A_1	610	610	—
	1305	1307	+0.2
	2920	2902	−0.6

TABLE X. *Methyl iodide*, CH_3I C−I, *2.10A*; CH_3 *as in ethane.*

	Observed	Calculated	Percent Error
E	885	882	−0.3
	1445	1470	+1.7
	3074	3020	−1.7
A_1	532	532	—
	1252	1253	+0.1
	2920	2902	−0.6

[15] In this connection, see J. H. Van Vleck and P. C. Cross, J. Chem. Phys. **1**, 357 (1933).

Some Mathematical Methods for the Study of Molecular Vibrations*

E. BRIGHT WILSON, JR.
Mallinckrodt Chemical Laboratory, Harvard University, Cambridge, Massachusetts

(Received October 19, 1940)

Developments which reduce the labor of calculating the vibration frequencies of complex molecules are described. In particular a vectorial scheme is given for obtaining the reciprocal of the matrix of the kinetic energy in terms of valence-type coordinates. A general rule for writing down the coefficients of the transformation to symmetry coordinates is derived together with a method of obtaining the kinetic energy reciprocal matrix ($\mathfrak{G}$) in terms of symmetry coordinates with a minimum of algebra. A treatment of redundant coordinates is developed. In addition, reduction of the secular equation by the splitting out of high frequencies, a new type of isotope product rule, and the determination of normal coordinates are discussed. The molecule CH_3Cl is worked out as an illustration.

IN a recent paper[1] a somewhat different mathematical machinery for the solution of the problem of small vibrations in molecules was put forward. Since then this system has been further developed and made easier to apply. It is the purpose of this paper to describe these developments.

VECTORIAL NOTATION

In I[1] it was found necessary to obtain the expressions

$$R_k=\sum_{i=1}^{3n}B_{ki}x_i, \tag{1}$$

in which R_k is an internal coordinate (valence-type coordinate) and x_i is one of the $3n$ Cartesian displacement coordinates of the n atoms. R_k might, for example, be the change in an interatomic distance or bond angle. It is convenient to describe the displacements of the atoms from their equilibrium configurations by using the vectors ϱ_t, one for each atom, instead of using the components x_i of these vectors. Then (1) becomes

$$R_k=\sum_{t=1}^{n}\mathbf{s}_{kt}\cdot\varrho_t, \tag{2}$$

in which the vectors $\mathbf{s}_{kt}$ have been introduced. The three B_{ki}'s belonging to atom t are the components of $\mathbf{s}_{kt}$. In this way a given internal coordinate R_k is characterized by a set of vectors $\mathbf{s}_{kt}$, one for each atom t. Actually these vanish except for those atoms whose motion changes R_k.

The kinetic energy matrix[2] G was shown in I to be given by

$$G_{kk'}=\sum_{i=1}^{3n}\mu_iB_{ki}B_{k'i}, \tag{3}$$

where μ_i is the reciprocal of the mass of the atom to which i refers. In terms of the vectors $\mathbf{s}_{kt}$, G is

$$G_{kk'}=\sum_{t=1}^{n}\mu_t\mathbf{s}_{kt}\cdot\mathbf{s}_{k't}. \tag{4}$$

This form has the advantage that it requires no coordinate system. Furthermore it is possible to express the vectors $\mathbf{s}_{kt}$ in terms of unit vectors directed along the chemical bonds, so that the problem of calculating G reduces to one of finding interbond angles. Moreover, very general expressions can be given for the vectors $\mathbf{s}_{kt}$ for the important types of internal coordinate R_k. For example if $\mathbf{r}_e$ is a *unit* vector from atom t to atom t' along the bond between them and R_k is the change in the length of this bond from its equilibrium value, then

$$\mathbf{s}_{kt}=-\mathbf{r}_e, \quad \mathbf{s}_{kt'}=+\mathbf{r}_e. \tag{5}$$

* Presented at the Symposium on the Structure of Molecules and Aggregates of Molecules at the Fifth Annual Symposium of the Division of Physical and Inorganic Chemistry of the American Chemical Society, Columbia University, New York, December 30, 1940 to January 1, 1941.

[1] E. B. Wilson, Jr., J. Chem. Phys. 7, 1047 (1939), hereafter called I.

[2] The kinetic energy is $2T=\dot{R}'G^{-1}\dot{R}$ or $2T=P'GP$, where P is a column matrix whose elements are the momenta conjugate to the R_k's. The prime denotes the transposed matrix.

If R_k is the change in the bond angle formed by the atoms $t'-t-t''$, t being at the apex, then the vectors are

$$\left.\begin{aligned}\mathbf{s}_{kt'} &= \epsilon(\cos\varphi\mathbf{r}_\epsilon-\mathbf{r}_\tau)/\sin\varphi,\\ \mathbf{s}_{kt''} &= \tau(\cos\varphi\mathbf{r}_\tau-\mathbf{r}_\epsilon)/\sin\varphi,\end{aligned}\right\}\ \text{(end atoms)}$$

$$\mathbf{s}_{kt} = [(\tau-\epsilon\cos\varphi)\mathbf{r}_\epsilon+(\epsilon-\tau\cos\varphi)\mathbf{r}_\tau]/\sin\varphi, \quad (6)$$

(apex atom)

in which the unit vectors $\mathbf{r}_\epsilon$ and $\mathbf{r}_\tau$ point out from the apex atom toward t' and t'', respectively, φ is the equilibrium value of the angle, ϵ and τ are the *reciprocals* of the bond lengths tt' and tt'', respectively. Similar expressions could be obtained for other types of internal coordinates.

Consequently, in order to obtain the kinetic energy matrix G, it is convenient to take the following steps. First indicate on a diagram of the molecule an arbitrary set of directions for the unit bond vectors. Calculate a table of dot products of these vectors. Write down the vector expressions (**s**) for the internal coordinates R_k in terms of these bond vectors. Compute the dot products of the vectors $\mathbf{s}_{kt}$ from these expressions and the table of bond dot products. Note, however, that for molecules with symmetry not all of these quantities may be needed, as discussed below.

Utilization of Symmetry

If the molecule contains sets of atoms which are symmetrically equivalent, for example, the carbon atoms in cyclopropane, it is unnecessary to deal with more than one atom in each set. What is desired in such cases is not the matrix G in terms of the internal coordinates R_k as described above but the related matrix $\mathfrak{G}$ in terms of internal symmetry coordinates $\mathfrak{R}_l$, since this leads to a secular equation factored because of the symmetry.[3] The coordinates R_k will form a representation (generally reducible) of the symmetry point group of the molecule so that linear combinations $\mathfrak{R}_l$ can be constructed which reduce the group. Let

$$\mathfrak{R}_l = \sum U_{lk}R_k, \quad (\text{or} \quad \mathfrak{R} = UR), \qquad (7)$$

where the transformation has been made orthogonal; i.e., $U^{-1}=U'$. (The prime denotes the transpose of the matrix.) Then the relation between G and $\mathfrak{G}$ is

$$\mathfrak{G} = UGU'. \qquad (8)$$

It is a consequence of the symmetry that the matrix $\mathfrak{G}$ will break up into diagonal blocks in the same way as the secular equation; that is, the blocks will correspond to the various irreducible representations of the symmetry group. Then it will be shown that for nondegenerate blocks

$$\mathfrak{G}_{ll'} = \sum_p \mu_p g_p \sum_{kk'} U_{lk}U_{l'k'}\mathbf{s}_{kt}\cdot\mathbf{s}_{k't}, \qquad (9)$$

in which g_p is the number of equivalent atoms in the pth set, while t is some one atom arbitrarily selected from those of the pth set. If the block corresponds to a degenerate irreducible representation, the result is

$$\mathfrak{G}_{ll'} = (1/d)\sum_p \mu_p g_p \sum_a^d \sum_{k,k'} U_{la,k}U_{l'a,k'}\mathbf{s}_{kt}\cdot\mathbf{s}_{k't}, \qquad (10)$$

in which the subscripts la, $l'a$ have been used instead of l, l', the a denoting the part of the degenerate representation meant. d is the degeneracy (2 or 3). The proof of this theorem will be found in Appendix I.

Equation (9) shows that it may be convenient to introduce the vectors

$$\mathbf{S}^{(t)}{}_l = \sum_k U_{lk}\mathbf{s}_{kt}, \qquad (11)$$

which bear the same relation to the symmetry coordinates $\mathfrak{R}_l$ as do the vectors $\mathbf{s}_{kt}$ to the internal coordinates R_k. In terms of the $\mathbf{S}$'s, $\mathfrak{G}$ becomes

$$\mathfrak{G}_{ll'} = (1/d)\sum_p \mu_p g_p \sum_a^d \mathbf{S}^{(t)}_{la}\cdot\mathbf{S}^{(t)}_{l'a}. \qquad (12)$$

Determination of Symmetry Coordinates

Ordinarily the coefficients U_{lk} have been determined largely by intuition and experience so as to give an orthonormal transformation with the desired symmetry properties. It is possible, however, to lay down some rules which make the choice of U's almost automatic.[4]

[3] See J. E. Rosenthal and G. M. Murphy, Rev. Mod. Phys. **8**, 317 (1936).

[4] See also E. Wigner, *Gruppentheorie* (Braunschweig, 1931).

First it will be assumed that each symmetry coordinate will be constructed from equivalent internal coordinates only.

Let

$$\mathfrak{R}_a=\sum_k U_{ak}R_k=(q/h)\sum_C U_{aC1}R_{C1}, \qquad (13)$$

where the sum over the internal coordinates k has been replaced by the sum over the operations C of the point group of the molecule, since these operations will permute the R's. The subscript $C1$ denotes the value of k obtained by operating C on an arbitrary coordinate R_1 of the set. q is the number of R's in the equivalent set used and h is the number of operations in the group, a factor necessary because the sum as written may include repetitions. If C' is some operation of the group, the result of operating on $\mathfrak{R}_a$ is

$$C'_{aa}\mathfrak{R}_a+C'_{ab}\mathfrak{R}_b=(q/h)\sum_C U_{aC1}R_{C'C1}, \qquad (14)$$

where C'_{aa} and C'_{ab} are the coefficients showing how the degenerate pair $\mathfrak{R}_a$ and $\mathfrak{R}_b$ transform under C' (if $\mathfrak{R}_a$ is nondegenerate, drop out $\mathfrak{R}_b$; if there is triple degeneracy, add another term).

Substituting Eq. (13) for $\mathfrak{R}_a$ and $\mathfrak{R}_b$, and equating coefficients of R_1, one gets

$$U_{aC1}=C_{aa}^{-1}U_{a1}+C_{ab}^{-1}U_{b1}=C_{aa}U_{a1}+C_{ba}U_{b1}, \qquad (15)$$

for the doubly-degenerate case, with similar results for the other cases. Furthermore, since the transformation U is orthonormal,[5]

$$1=\sum_k U_{ak}^2=(q/h)\sum_C U_{aC1}^2$$

$$=(q/h)\sum_C(C_{aa}U_{a1}+C_{ba}U_{b1})^2=(q/2)(U_{a1}^2+U_{b1}^2).$$

Here the group theory result[3] has been used that

$$\sum_C C_{aa}C_{aa}=h/d, \quad \sum_C C_{aa}C_{ba}=0, \quad \text{etc.}, \qquad (16)$$

d being the degeneracy of the irreducible representation involved. The final conclusions are that

$$U_{ak}=\pm(1/q)^{\frac{1}{2}},$$
$$U_{ak}^2+U_{bk}^2=2/q,$$
$$U_{ak}^2+U_{bk}^2+U_{ck}^2=3/q \qquad (17)$$

for degeneracies 1, 2, and 3, respectively.

These rules determine the nondegenerate coefficients completely except for a constant multiplier ±1, which is unimportant. In the degenerate cases, if $q=h$, there will be *d sets* $\mathfrak{R}_a, \mathfrak{R}_b, \cdots$ in a given irreducible representation so that the ambiguity inherent in Eq. (17) is needed in order to provide for the several sets. So long as these are made orthogonal to one another the choice is arbitrary. If $q<h$ further restrictions are placed on the U's. Thus, from Eq. (13)

$$U_{a',k'}=(q/h)\sum_{C_{k'},a''} C_{a''a'}U_{a'',k},$$

where the sum is over all operations $C_{k'}$ such that $C_{k'}R_k\rightarrow R_{k'}$. In particular if $k'=k$, the operations $C_{k'}$ form a subgroup of order h/q. Then *if* $\mathfrak{R}_a$, $\mathfrak{R}_b$, etc. reduce this subgroup ($\mathbf{C}_k$) as well as $\mathbf{C}$, then

$$\sum_{C_k} C_{a''a'}=\sum_{C_k} C^{\Gamma}_{a''a'}C^{A_1}_{aa}=\delta\Gamma_{A_1}\delta_{a''a}\delta_{a'a}(h/q) \qquad (18)$$

will vanish unless $a''=a'$ and unless $\mathfrak{R}_{a'}$ falls in the A_1 representation of $\mathbf{C}_k$. ($C^{A_1}_{aa}=1$.) In this case therefore $U_{la',k}$ vanishes unless $\mathfrak{R}_{a'}$ is in A_1 of the subgroup. In many important cases only one of $\mathfrak{R}_a$, $\mathfrak{R}_b$, etc. will fall in A_1 so that only one of $U_{la,k}$, $U_{lb,k}$, etc. will not vanish.

If the "orientation" of $\mathfrak{R}_a$, $\mathfrak{R}_b$, etc. is such that they do not reduce the subgroup $\mathbf{C}_k$, then the transformation coefficients $C_{a''a'}$ will be related to coefficients $C_{AA'}$ in terms of coordinates which do reduce $\mathbf{C}_k$ by the equation

$$C_{a''a'}=\sum_{A'A''} L_{A'a'}L_{A''a'}C_{A'A''}$$

so that

$$\sum_{C_k} C_{a''a'}=\sum_{C_k} C_{a''a'}C^{A_1}_{AA}$$

$$=\sum_{A'A''} L_{A'a''}L_{A''a'}\sum_{C_k} C^{\Gamma}_{A'A''}C^{A_1}_{AA}$$

$$=\delta\Gamma_{A_1}L_{Aa''}L_{Aa'}(h/g).$$

$$\therefore \quad U_{la',k}=\delta\Gamma_{A_1}L_{Aa'}\sum_{a''} L_{Aa''}U_{la'',k}, \qquad (19)$$

[5] Only real coordinates will be treated in this paper. While there is probably no difficulty in handling complex coordinates, they are quite unnecessary for practical work, since no molecule of symmetry C_n, C_{nh}, or S_n ($n>2$) has yet been studied, or is likely to be.

so if Γ (the representation of $\mathbf{C}_k$ in which $C_{aa'}$ occurs) is A_1,

$$L_{Ab}U_{la,k}=L_{Aa}U_{lb,k},$$

for example. (This result applies only if the reduced representation of $\mathfrak{R}_a$, $\mathfrak{R}_b$, $\cdots$ has one and only one combination in A_1 of $\mathbf{C}_k$.)

The considerations of this section are much more elaborate than are necessary for most simple molecules, where intuition yields the correct result on inspection, but may be helpful in very complicated cases.

A Further Reduction

The methods described in the previous sections are quite satisfactory in many cases and do not involve much labor. However, in molecules of high symmetry containing atoms in special positions it is possible to carry out a further reduction.

As before, select one typical atom from each set of equivalent atoms. One or more of the operations of the molecular point group $\mathbf{C}$ will leave such an atom unshifted. These operations will form a subgroup $\mathbf{C}_t$ of the whole group $\mathbf{C}$. Furthermore the internal coordinates R_k which involve the atom t will divide into sets equivalent under $\mathbf{C}_t$. Note that R's equivalent under $\mathbf{C}$ may sometimes not be equivalent under $\mathbf{C}_t$. Now select one typical coordinate R_{ki} from each set i of equivalent (under $\mathbf{C}_t$) R's. Obtain its vector $\mathbf{s}_{kit}$. Orient a set of Cartesian axes x, y, z so that these coordinates form the basis of a fully reduced representation of both $\mathbf{C}_t$ and $\mathbf{C}$. This is easily done by inspection. If there is any ambiguity make the y component of $\mathbf{s}_{kit}$ zero for some choice of ki. Then write down the x, y, and z components of the typical vectors $\mathbf{s}_{kit}$, one for each set i.

The elements of the kinetic energy matrix $\mathfrak{G}$ may [see Eq. (12)] be written as

$$\mathfrak{G}_{ll'}=(1/d)\sum_p \mu_p g_p \sum_{a'=1}^{d}\sum_{i,j} \mathfrak{B}^{(t)}_{la',i}\cdot\mathfrak{B}^{(t)}_{l'a',j}, \quad (20)$$

where

$$\mathfrak{B}^{(t)}_{la',i}=\sum_{k(\text{in } i)} U_{la',k}\mathbf{s}_{kit}. \quad (21)$$

But

$$\mathfrak{B}^{(t)}_{la',i}\cdot\mathfrak{B}^{(t)}_{l'a',j}=(\mathfrak{B}^{(t)}_{la',i}\cdot\mathbf{a})(\mathfrak{B}^{(t)}_{l'a',j}\cdot\mathbf{a}) + (\mathfrak{B}^{(t)}_{la',i}\cdot\mathbf{b})(\mathfrak{B}^{(t)}_{l'a',j}\cdot\mathbf{b})+(\mathfrak{B}^{(t)}_{la',i}\cdot\mathbf{c})(\mathfrak{B}^{(t)}_{l'a',j}\cdot\mathbf{c}), \quad (22)$$

where a, b, and c are unit vectors in the x, y, z directions.

Since $\mathfrak{B}^{(t)}_{la',i}$ is a vector it must fall into an irreducible representation of $\mathbf{C}_t$ containing one or more of x, y and z, unless it is zero. If C_t is of high symmetry, x, y and z will be in different irreducible representations or in different parts of a degenerate representation. In this case only one dot product of the type $\mathfrak{B}^{(t)}_{la',i}\cdot\mathbf{a}$ will be nonvanishing for given a'. For lower symmetries, however, two or more of x, y, z may fall in some irreducible representation of $\mathbf{C}_t$.

A dot product $\mathfrak{B}^{(t)}_{la',i}\cdot\mathbf{a}''$ can be considerably reduced as follows.

$$\begin{aligned}\mathfrak{B}^{(t)}_{la',i}\cdot\mathbf{a}''&=\sum_{k(\text{in } i)} U_{la',k}\mathbf{s}_{kit}\cdot\mathbf{a}''\\ &=(q_i/h_t)\sum_{C_t} U_{la',Ck}(\mathbf{s}_{Ck,it}\cdot\mathbf{a}'')\\ &=(q_i/h_t)\sum_{C_t}\sum_{a'''} C_{a'''a'}U_{la''',k}(\mathbf{s}_{Ck,it}\cdot\mathbf{a}''),\end{aligned} \quad (23)$$

in which q_i is the number of coordinates R_k equivalent under $\mathbf{C}_t$, h_t is the order of $\mathbf{C}_t$. The sum over a''' means the sum over the components a, b, c (or one, or two of them) depending upon the degeneracy (under $\mathbf{C}_t$) of the representation in which $\mathfrak{B}^{(t)}_{la',i}$ occurs. $C_{a'''a'}$ is a matrix element of the transformation matrix for $\mathfrak{R}_{la'}$ representing the symmetry operation C. See Eq. (15). A dot product of two vectors is unchanged if both are operated on by a symmetry operation C^{-1} so

$$\mathbf{s}_{Ck,it}\cdot\mathbf{a}''=\mathbf{s}_{kit}\cdot\sum_{a^{\text{iv}}} C_{a^{\text{iv}}a''}\mathbf{a}^{\text{iv}} \quad (24)$$

and therefore

$$\mathfrak{B}^{(t)}_{la',i}\cdot\mathbf{a}''=\psi_{a'a''}(q_i/d_{li})\sum_{a'''}U_{la''',k}(\mathbf{s}_{kit}\cdot\mathbf{a}'''), \quad (25)$$

in which $\psi_{a'a''}$ is unity if $a''=a'$ or if a'' is in the same nondegenerate representation of $\mathbf{C}_t$ as a', but $\psi_{a'a''}=0$ if a' and a'' are in different parts of the same degenerate representation or in different representations. The sum over a''' means the sum over the degenerate components (a, b, c) of which a' is a member (under $\mathbf{C}_t$!), provided that a' is a member of a degenerate set. Equation (16) has been used. d_{li} is the degeneracy under $\mathbf{C}_t$.

Therefore only one representative vector $\mathbf{s}_{kil}$ from each set i for each typical atom t need be found. Furthermore, the coefficients $U_{la'',k}$ can be obtained from the procedure of the previous section and only certain of them are needed.

This method is illustrated later for CH_3Cl.

Redundant Coordinates

It frequently happens that, if complete sets of symmetrically equivalent internal coordinates are chosen, there are more coordinates than are independent. For example, there are six H−C−H angles in CH_4, all equivalent, but there exists one relation between them so that they are not all independent. It is clearly not good to pick out five arbitrarily for use; such a method spoils the symmetry. There are several schemes which may be used.

The simplest is to note that if all the internal coordinates of a given set are treated as forming a reducible representation of the symmetry group of the molecule and if this representation is reduced, the linear combinations which represent conditions on the coordinates will turn up among the irreducible representations and can be recognized as such. Thus in CH_4, the sum of the six angles, which must be a constant, appears in A_1 of the point group $\mathbf{T}_d$. These redundant combinations are then merely dropped. In other words, instead of arbitrarily dropping one or more of the internal coordinates R_k, certain of the symmetry coordinates $\Re_l$ are ignored. This is justifiable because these coordinates are identically zero and therefore disappear from the expressions for the potential and kinetic energies.

In many cases the conditions on the coordinates involve more than one kind of internal coordinate R_k. Thus in CH_3Cl the sum of the three H−C−H angles (α_{ij}) plus the sum of the three H−C−Cl angles (β_i) must be constant. It is usually most convenient to construct the symmetry coordinates $\Re_l$ out of internal coordinates of one kind. In the example of CH_3Cl there are essentially three possible methods of proceeding. In the first, the condition on the angles is made one of the symmetry coordinates and the other symmetry coordinates are made orthogonal to it. The "condition coordinate" is then ignored. This process can be carried out either in one or in two steps. The second method is to make up the symmetry coordinates in the usual way out of the α's and the β's separately. Then the coordinate $\alpha_{12}+\alpha_{23}+\alpha_{31}$ is ignored. The third way is to ignore $\beta_1+\beta_2+\beta_3$.

For a given molecule, any of these methods is legitimate, but they do not correspond to the same choice of potential function. Thus if the potential energy is written as

$$\begin{aligned}2V=&F_\alpha(\alpha_{12}^2+\alpha_{23}^2+\alpha_{31}^2)+F_\beta(\beta_1^2+\beta_2^2+\beta_3^2)\\&+2F_{\alpha\alpha}(\alpha_{12}\alpha_{23}+\alpha_{23}\alpha_{31}+\alpha_{31}\alpha_{12})\\&+2F_{\beta\beta}(\beta_1\beta_2+\beta_2\beta_3+\beta_3\beta_1)\\&+2F_{\alpha\beta}(\beta_1\alpha_{12}+\beta_2\alpha_{23}+\beta_3\alpha_{31}\\&+\beta_2\alpha_{12}+\beta_3\alpha_{23}+\beta_1\alpha_{31})\\&+2F'_{\alpha\beta}(\beta_1\alpha_{23}+\beta_2\alpha_{31}+\beta_3\alpha_{12}),\end{aligned}\tag{26}$$

only four of the six constants F are independent and they are therefore not fully defined without further information. It is therefore only possible to evaluate four linear combinations of the six F's. The α's and β's can be combined into the symmetry coordinates

$$\begin{aligned}\Re_{\alpha A}&=(1/\sqrt{3})(\alpha_{12}+\alpha_{23}+\alpha_{31}),\\\Re_{\alpha Ea}&=(1/\sqrt{2})(\alpha_{12}-\alpha_{23}),\\\Re_{\alpha Eb}&=(1/\sqrt{6})(\alpha_{12}+\alpha_{23}-2\alpha_{31}),\end{aligned}\tag{27}$$

$$\begin{aligned}\Re_{\beta A}&=(1/\sqrt{3})(\beta_1+\beta_2+\beta_3),\\\Re_{\beta Ea}&=(1/\sqrt{2})(\beta_1-\beta_3),\\\Re_{\beta Eb}&=(1/\sqrt{6})(2\beta_2-\beta_1-\beta_3).\end{aligned}\tag{28}$$

In terms of these the potential energy becomes

$$\begin{aligned}2V=&\mathfrak{F}_{\alpha A}\Re_{\alpha A}^2+\mathfrak{F}_{\beta A}\Re_{\beta A}^2+2\mathfrak{F}_{\alpha\beta A}\Re_{\alpha A}\Re_{\beta A}\\&+\mathfrak{F}_{\alpha E}(\Re_{\alpha Ea}^2+\Re_{\alpha Eb}^2)+\mathfrak{F}_{\beta E}(\Re_{\beta Ea}^2+\Re_{\beta Eb}^2)\\&+2\mathfrak{F}_{\alpha\beta E}(\Re_{\alpha Ea}\Re_{\beta Ea}+\Re_{\alpha Eb}\Re_{\beta Eb}),\end{aligned}\tag{29}$$

in which $\mathfrak{F}_{\alpha A}=F_\alpha+2F_{\alpha\alpha}$, $\mathfrak{F}_{\beta A}=F_\beta+2F_{\beta\beta}$, $\mathfrak{F}_{\alpha\beta A}=2F_{\alpha\beta}+F'_{\alpha\beta}$, $\mathfrak{F}_{\alpha E}=F_\alpha-F_{\alpha\alpha}$, $\mathfrak{F}_{\beta E}=F_\beta-F_{\beta\beta}$, $\mathfrak{F}_{\alpha\beta E}=F_{\alpha\beta}-F'_{\alpha\beta}$.

If the third method is used, i.e., ignoring $\Re_{\beta A}$, the potential function employed is (29) with $\mathfrak{F}_{\beta A}=\mathfrak{F}_{\alpha\beta A}=0$, or $F_\beta+2F_{\beta\beta}=0$, $2F_{\alpha\beta}+F'_{\alpha\beta}=0$. The second method corresponds to putting $\mathfrak{F}_{\alpha A}$ and $\mathfrak{F}_{\alpha\beta A}$ equal to zero. The first method, on the other hand, requires the further transformation

$$\begin{aligned}\Re_{1A}&=(1/\sqrt{2})(\Re_{\alpha A}+\Re_{\beta A}),\\\Re_{2A}&=(1/\sqrt{2})(\Re_{\alpha A}-\Re_{\beta A}),\end{aligned}\tag{30}$$

$\Re_{1A}=0$ being the condition on the coordinates,

so that $\mathfrak{R}_{1A}$ is ignored. The potential energy (A terms) is

$$\mathfrak{F}_{11}\mathfrak{R}_{1A}^2+\mathfrak{F}_{22}\mathfrak{R}_{2A}^2+2\mathfrak{F}_{12}\mathfrak{R}_{1A}\mathfrak{R}_{2A}, \qquad (31)$$

with $\mathfrak{F}_{11}=\frac{1}{2}(\mathfrak{F}_{\alpha A}+\mathfrak{F}_{\beta A}+2\mathfrak{F}_{\alpha\beta A})$, $\mathfrak{F}_{22}=\frac{1}{2}(\mathfrak{F}_{\alpha A}+\mathfrak{F}_{\beta A}-2\mathfrak{F}_{\alpha\beta A})$, $\mathfrak{F}_{12}=\frac{1}{2}(\mathfrak{F}_{\alpha A}-\mathfrak{F}_{\beta A})$. Thus ignoring $\mathfrak{R}_{1A}$ is equivalent to putting $\mathfrak{F}_{11}=\mathfrak{F}_{12}=0$, or $\mathfrak{F}_{\alpha A}=\mathfrak{F}_{\beta A}$, $\mathfrak{F}_{\alpha\beta A}=-\mathfrak{F}_{\alpha A}$. The combinations of the original six force constants which can be evaluated are thus: $\mathfrak{F}_{22}=\frac{1}{2}(F_\alpha+F_\beta+2F_{\alpha\alpha}+2F_{\beta\beta}-4F_{\alpha\beta}-2F'_{\alpha\beta})$, besides $\mathfrak{F}_{\alpha E}$, $\mathfrak{F}_{\beta E}$, and $\mathfrak{F}_{\alpha\beta E}$.

If the most general potential function is used, it makes no difference which of the above systems is adapted as far as a given molecule is concerned. However, if force constants for molecules of different symmetries are being compared (for example, CH_4 and CH_3D) it is necessary that they be reduced to the same basis. This is easily done. Thus, using the third method, $\mathfrak{F}_{\alpha A}$ is obtained and $\mathfrak{F}_{\beta A}$ and $\mathfrak{F}_{\alpha\beta A}$ are set equal to zero. But $\mathfrak{F}_{22}=\frac{1}{2}(\mathfrak{F}_{\alpha A}+\mathfrak{F}_{\beta A}-2\mathfrak{F}_{\alpha\beta A})$ so the force constant $\mathfrak{F}_{22}$ obtained by the first method equals one-half the constant $\mathfrak{F}_{\beta A}$ obtained by the third method.

Similar methods can be applied to other sets of force constants and to other, more complicated molecules. A condition on a set of coordinates can be recognized because the corresponding vector $\mathfrak{S}_t^{(i)}$ will be zero for every atom t.

Splitting of High Frequencies

It has been pointed out in I that when one or more frequencies occur which are considerably higher than the rest, it is possible to reduce the degree of the secular equation by putting certain force constants equal to infinity when solving for the low frequencies. Any hydrocarbon provides an illustration since the C−H stretching frequencies are considerably higher than the rest. If the C−H force constants are put equal to infinity, the other frequencies can be obtained with less than one percent error.

It is shown in Appendix II that if $\mathfrak{G}$ is the kinetic energy matrix and, if this is divided into submatrices:

$$\begin{pmatrix}\mathfrak{G}_{11} & \mathfrak{G}_{12}\\ \mathfrak{G}_{21} & \mathfrak{G}_{22}\end{pmatrix},$$

where the rows and columns denoted by 1 represent high frequency coordinates, while 2 corresponds to the rest of the coordinates, then the smaller matrix

$$\mathfrak{G}^0=\mathfrak{G}_{22}-\mathfrak{G}_{21}\mathfrak{G}_{11}^{-1}\mathfrak{G}_{12} \qquad (32)$$

can be used with the force constant matrix $\mathfrak{F}_{22}$ in the usual way to construct the smaller secular equation for the low frequencies. Here $\mathfrak{F}_{22}$ is the force constant matrix with high frequency rows and columns omitted.

This technique is particularly simple if there is only one high frequency in a given factor. In this case (32) becomes

$$\mathfrak{G}^0_{ll'}=\mathfrak{G}_{ll'}-\mathfrak{G}_{l1}\mathfrak{G}_{1l'}/\mathfrak{G}_{11}, \qquad (33)$$

where the quantities involved are numbers, not submatrices.

Isotope Rule

Teller[6] and Redlich[7] have shown that the product of the frequencies of a given molecule divided by the similar product for the same molecule with some of its atoms replaced by isotopic atoms is independent of the potential function and depends only on the masses and geometry. This can be shown to be true also by the use of the matrix $\mathfrak{G}$, and the expression for the frequency quotient can be obtained in a somewhat different form. From I[1] it is seen that the constant term in the secular equation (or in any factor of it if it is factored) is given by

$$|\mathfrak{G}||\mathfrak{F}|$$

where $|\mathfrak{G}|$ is the determinant of the matrix $\mathfrak{G}$ and $|\mathfrak{F}|$ is the determinant of the force constant matrix. But the constant term is the product of the roots of the equation. Furthermore the force constants are the same for the two isotopic species so that

$$\frac{\prod^{\lambda}\lambda_\lambda}{\prod^{\lambda}\lambda_\lambda'}=\frac{|\mathfrak{G}|}{|\mathfrak{G}'|}, \qquad (34)$$

where $\lambda_\lambda=4\pi^2\nu_\lambda^2$, ν_λ being a frequency. The prime denotes the isotopic species. The product contains all λ's of the same symmetry as the factor $\mathfrak{G}$ used.

[6] Cf. C. K. Ingold *et al.*, J. Chem. Soc. London, 971 (1936).

[7] O. Redlich, Zeits. f. physik. Chemie **B28**, 371 (1935).

A useful approximation which is new is to apply the same equation to the low frequencies only, using the reduced matrix $\mathfrak{G}^0$ instead of $\mathfrak{G}$. This enables the product rule test to be applied separately to the high and to the low frequencies, to a good approximation.

Determination of Normal Coordinates

In terms of the coordinates R the potential and kinetic energies are

$$2V=R'FR, \quad 2T=\dot{R}'G^{-1}\dot{R}. \tag{35}$$

The normal coordinates Q are linearly related to the coordinates R by a transformation

$$R=LQ \tag{36}$$

in which L is chosen so that the energies in terms of Q have the forms

$$\begin{aligned} 2V&=Q'L'FLQ=Q'\Lambda Q, \\ 2T&=\dot{Q}'L'G^{-1}L\dot{Q}=\dot{Q}'E\dot{Q}, \end{aligned} \tag{37}$$

in which Λ is a diagonal matrix (whose diagonal elements will be the λ's), while E is the unit matrix. Therefore

$$L'FL=\Lambda, \quad L'G^{-1}L=E \tag{38}$$

or

$$GFL=L\Lambda, \tag{39}$$

the latter being the set of simultaneous equations which determine the transformation L. Similarly,

$$L^{-1}GF=\Lambda L^{-1}. \tag{40}$$

Written out, these are

$$\sum_l[(GF)_{kl}-\delta_{kl}\lambda_\lambda]L_{l\lambda}=0, \quad (k=1, 2, 3, \cdots), \tag{41}$$

$$\sum_k[(GF)_{kl}-\delta_{kl}\lambda_\lambda]L^{-1}_{\lambda k}=0, \quad (l=1, 2, 3, \cdots), \tag{42}$$

where $\delta_{kl}=1$ if $k=l$, zero otherwise. Thus the elements in the rows of the secular equation provide the coefficients in the forward transformation (L), while the columns serve for the backward transformation (L^{-1}).

These homogeneous equations give only the ratios of the $L_{l\lambda}$ or $L^{-1}_{\lambda k}$ for a given λ, so that it is usually necessary to normalize them. The normalization condition is that $LL'=G$ or $L^{-1}G(L')^{-1}=E$. The second of these is easy to apply. Let $(L_0^{-1})_{\lambda k}$ be an unnormalized value and K_λ the normalizing constant. Then

$$K_\lambda=1/[\sum_{kl}(L_0^{-1})_{\lambda k}(L_0^{-1})_{\lambda l}G_{kl}]^{\frac{1}{2}}. \tag{43}$$

Unfortunately, it is not easy to normalize L, it being necessary to solve the simultaneous equations

$$\sum_\lambda K_\lambda^2(L^0_{k\lambda})^2=G_{kk}. \tag{44}$$

Probably a better way to get the transformation L, in case the normalized values are required, is to solve first for L^{-1}, carry out the normalization as indicated above, then use the relation

$$L'=L^{-1}G. \tag{45}$$

Naturally if the molecule has symmetry, $\mathfrak{G}$, $\mathfrak{F}$ and $\mathfrak{R}$ are used instead of G, F and R.

An Example: CH_3Cl

Choose the changes in the three C−H distances (r_1, r_2, r_3), the C−Cl distance (R), the three H−C−H angles (α_{12}, α_{23}, α_{31}), and the three H−C−Cl angles (β_1, β_2, β_3) as the internal coordinates. One of these 10 is redundant since there are only $3\times4-6=9$ internal degrees of freedom The point group **C** for the whole molecule is $\mathbf{C}_{3v}$. The sets of equivalent atoms and the subgroup for one of each set are H_1, H_2, $H_3(\mathbf{C}_s)$; $C(\mathbf{C}_{3v})$; $Cl(\mathbf{C}_{3v})$. The usual group theory arguments show that the irreducible representations A_1, A_2, E of $\mathbf{C}_{3v}$ will contain the following combinations of the internal coordinates. A_1: r, R, α, (β), E: r, α, β. The combination of β's found in A_1 will be considered the redundant coordinate and ignored. The secular equation will thus factor into three cubics, of which two are identical. The $\mathfrak{G}$ matrix will factor in the same way.

The atoms C and Cl have the symmetry of the whole molecule and therefore their subgroups need no discussion. H_1 has only the identity E and a plane of symmetry σ. The coordinates touching H_1 are r_1, β_1, α_{12}, α_{31}. The latter two form a set equivalent under this group. The two representations of $\mathbf{C}_s$ contain: A: r_1, β_1, one combination of α's, B: one combination of α's. The coordinates x, y, z have the symmetries: $\mathbf{C}_{3v}$: $A_1(z)$, E (x, y degenerate pair); $\mathbf{C}_s$: A (x, z nondegenerate), $B(y)$.

For the Cl atom the vectors needed are:

$$\mathbf{s}_R=\mathbf{R},\quad \mathbf{s}_{\beta_1}=-(\tfrac{1}{4})\sqrt{2}\tau(\mathbf{R}+3\mathbf{r}_1),$$

in which $\mathbf{R}$ and $\mathbf{r}_1$ are unit vectors from the carbon along the bonds to Cl and H_1, respectively, and τ is the reciprocal of the equilibrium C−Cl length. Tetrahedral angles have been used. The nonvanishing components of these vectors are:

$$\mathbf{s}_R\cdot\mathbf{c}=+1;\quad \mathbf{s}_{\beta_1}\cdot\mathbf{a}=-\tau.$$

The symmetry coordinates involving the Cl atom are R_{A_1} and β_{E_a}, β_{E_b}. This notation tells which set of equivalent internal coordinates are combined and to which representation of $\mathbf{C}_{3v}$ it belongs. The following U's are needed and are obtained by the methods already explained. For R, $q=1$; for β, $q=3$; $d=1$ for A_1, $d=2$ for E.

$$U_{R(A_1),R}=1,\quad U_{\beta(E_a),\beta_1}=\tfrac{2}{3}\tfrac{1}{2},\quad U_{\beta(E_b),\beta_1}=0.$$

Consequently, the only nonvanishing components are

$$\overset{(\mathrm{Cl})}{\mathfrak{S}}_{R(A_1)}\cdot\mathbf{c}=+1,\quad \overset{(\mathrm{Cl})}{\mathfrak{S}}_{\beta(E_b)}\cdot\mathbf{b}=-\tfrac{3}{2}\tfrac{1}{2}\tau,$$
$$\overset{(\mathrm{CL})}{\mathfrak{S}}_{\beta(E_a)}\cdot\mathbf{a}=-\tfrac{3}{2}\tfrac{1}{2}\tau.$$

The contributions of the Cl atom to the kinetic energy matrix $\mathfrak{G}$ are thus (since $g_{\mathrm{Cl}}=1$):

$$\text{to } \mathfrak{G}_{R(A_1)R(A_1)}:\mu_{\mathrm{Cl}};\quad \text{to } \mathfrak{G}_{\beta(E_a),\beta(E_a)}:\tfrac{3}{2}\tau^2\mu_{\mathrm{Cl}}.$$

The treatment of the C atom is so similar that it will not be given except to mention that the projections on the xy plane of the vectors for β_1 and α_{12} make an angle of 60° with one another, a fact which must be taken into account in obtaining the coefficients U.

For H atom 1 the vectors needed are

$$\mathbf{s}_{r_1}=\mathbf{r}_1;\quad \mathbf{s}_{\beta_1}=-\tfrac{1}{4}\sqrt{2}\epsilon(\mathbf{r}_1+3\mathbf{R}),$$
$$\mathbf{s}_{\alpha_{12}}=-\tfrac{1}{4}\sqrt{2}\epsilon(\mathbf{r}_1+3\mathbf{r}_2).$$

Here ϵ is the reciprocal of the C−H distance. The nonvanishing components are:

$$\mathbf{s}_{r_1}\cdot\mathbf{a}=\tfrac{2}{3}\sqrt{2},\quad \mathbf{s}_{r_1}\cdot\mathbf{c}=-\tfrac{1}{3},$$
$$\mathbf{s}_{\beta_1}\cdot\mathbf{a}=-\tfrac{1}{3}\epsilon,\quad \mathbf{s}_{\beta_1}\cdot\mathbf{c}=-\tfrac{2}{3}\sqrt{2}\epsilon,$$
$$\mathbf{s}_{\alpha_{12}}\cdot\mathbf{a}=\tfrac{1}{6}\epsilon,\quad \mathbf{s}_{\alpha_{12}}\cdot\mathbf{b}=-\tfrac{1}{2}\sqrt{3}\epsilon,\quad \mathbf{s}_{\alpha_{12}}\cdot\mathbf{c}=\tfrac{1}{3}\sqrt{2}\epsilon.$$

The U's needed are given below. For r, and β, $q=3$ but $q_i=1$; for α, $q=3$, $q_i=2$. $d_i=1$. The symmetry coordinates involved are r_{A_1}, r_{E_a}, r_{E_b}, β_{E_a}, β_{E_b}, α_{A_1}, α_{E_a}, α_{E_b}.

$$U_{r(A_1),r_1}=\tfrac{1}{3}\tfrac{1}{2};\quad U_{\alpha(A_1),\alpha_{12}}=\tfrac{1}{3}\tfrac{1}{2};$$
$$U_{r(E_a),r_1}=\tfrac{2}{3}\tfrac{1}{2};\quad U_{r(E_b),r_1}=0;$$
$$U_{\beta(E_a),\beta_1}=\tfrac{2}{3}\tfrac{1}{2};\quad U_{\beta(E_b),\beta_1}=0;$$
$$U_{\alpha(E_a),\alpha_{12}}=\tfrac{1}{6}\tfrac{1}{2};\quad U_{\alpha(E_b),\alpha_{12}}=\tfrac{1}{2}\tfrac{1}{2}.$$

The nonvanishing components needed are:

$$\overset{(H_1)}{\mathfrak{S}}_{r(A_1)}\cdot\mathbf{a}=\tfrac{2}{9}\sqrt{6};\quad \overset{(H_1)}{\mathfrak{S}}_{r(A_1)}\cdot\mathbf{c}=-\tfrac{1}{9}\sqrt{3};$$
$$\overset{(H_1)}{\mathfrak{S}}_{\alpha(A_1)}\cdot\mathbf{a}=\tfrac{1}{9}\sqrt{3}\epsilon;\quad \overset{(H_1)}{\mathfrak{S}}_{\alpha(A_1)}\cdot\mathbf{c}=\tfrac{2}{9}\sqrt{6}\epsilon;$$
$$\overset{(H_1)}{\mathfrak{S}}_{r(E_a)}\cdot\mathbf{a}=\tfrac{4}{9}\sqrt{3};\quad \overset{(H_1)}{\mathfrak{S}}_{r(E_a)}\cdot\mathbf{c}=-\tfrac{1}{9}\sqrt{6};$$
$$\overset{(H_1)}{\mathfrak{S}}_{\beta(E_a)}\cdot\mathbf{a}=-\tfrac{1}{9}\sqrt{6}\epsilon;\quad \overset{(H_1)}{\mathfrak{S}}_{\beta(E_a)}\cdot\mathbf{c}=-\tfrac{4}{9}\sqrt{3}\epsilon,$$
$$\overset{(H_1)}{\mathfrak{S}}_{\alpha(E_a)}\cdot\mathbf{a}=\tfrac{1}{18}\sqrt{6}\epsilon;\quad \overset{(H_1)}{\mathfrak{S}}_{\alpha(E_a)}\cdot\mathbf{c}=\tfrac{2}{9}\sqrt{3}\epsilon;$$
$$\overset{(H_1)}{\mathfrak{S}}_{\alpha(E_b)}\cdot\mathbf{b}=-\tfrac{1}{2}\sqrt{6}\epsilon.$$

These contribute the terms in μ_{H} found in the complete $\mathfrak{G}$ matrices, which are given below.

A_1	r	α	R
r	$\mu_{\mathrm{H}}+\tfrac{1}{3}\mu_{\mathrm{C}}$	$-\tfrac{2}{3}\sqrt{2}\epsilon\mu_{\mathrm{C}}$	$-\tfrac{1}{3}\sqrt{3}\mu_{\mathrm{C}}$
α		$\epsilon^2\mu_{\mathrm{H}}+\tfrac{8}{3}\epsilon^2\mu_{\mathrm{C}}$	$\tfrac{2}{3}\sqrt{6}\epsilon\mu_{\mathrm{C}}$
R			$\mu_{\mathrm{C}}+\mu_{\mathrm{Cl}}$

E	r	α	β
r	$\mu_{\mathrm{H}}+\tfrac{4}{3}\mu_{\mathrm{C}}$	$\tfrac{4}{3}\sqrt{2}\epsilon\mu_{\mathrm{C}}$	$\sqrt{2}(\tau+\tfrac{1}{3}\epsilon)\mu_{\mathrm{C}}$
α		$\tfrac{5}{2}\epsilon^2\mu_{\mathrm{H}}+\tfrac{8}{3}\epsilon^2\mu_{\mathrm{C}}$	$-\tfrac{1}{2}\epsilon^2\mu_{\mathrm{H}}+2\epsilon(\tau+\tfrac{1}{3}\epsilon)\mu_{\mathrm{C}}$
β			$\epsilon^2\mu_{\mathrm{H}}+\tfrac{3}{2}(\tau+\tfrac{1}{3}\epsilon)^2\mu_{\mathrm{C}}+\tfrac{3}{2}\tau^2\mu_{\mathrm{Cl}}$.

Since $\mathfrak{G}$ is symmetrical, only half the elements are given. The secular equation can be obtained from $\mathfrak{G}$ either in expanded form as shown in I or as a determinantal equation $|\mathfrak{G}\mathfrak{F}-E\lambda|=0$. In the present example it is very convenient to split out the high C−H frequencies by the technique already described.

It should be noted that any procedure such as this must be used very carefully if mistakes are to be avoided. It is in addition desirable to find the elements of $\mathfrak{G}$ by several different methods in order to insure that no mistake has been made.

Appendix I

We wish to prove that

$$\sum_{t\ (\text{in } p)}^{g_p} \mathfrak{S}_{la'}^{(t)} \cdot \mathfrak{S}_{l'a'}^{(t)} = (g_p/d) \sum_{a''}^{d} \mathfrak{S}_{la''}^{(t')} \cdot \mathfrak{S}_{l'a''}^{(t')},$$

in which the first sum is over all the g_p equivalent atoms of a set p, the second sum is over the d degenerate components. t' is an arbitrary atom of the set p.

Denote by Ct' the atom which originally occupied the position to which t moves when the symmetry operation C is applied to the molecule. Then if h is the number of operations in the group,

$$\sum_{t}^{g_p} \mathfrak{S}_{la'}^{(t)} \cdot \mathfrak{S}_{l'a'}^{(t)} = (g_p/h) \sum_{C} \mathfrak{S}_{la'}^{(Ct')} \cdot \mathfrak{S}_{l'a'}^{(Ct')},$$

the factor (g_p/h) taking account of duplications. If a symmetry operation is applied to the two vectors $\mathfrak{S}_{la'}^{(t')}$ and $\mathfrak{S}_{l'a'}^{(t')}$ their dot product will not be altered. The transformation may be written

$$\mathfrak{S}_{la'}^{(Ct')} \cdot \mathfrak{S}_{l'a'}^{(Ct')} = (C_{a'a}^{-1} \mathfrak{S}_{la}^{(t')} + C_{a'b}^{-1} \mathfrak{S}_{lb}^{(t')}) \cdot (C_{a'a}^{-1} \mathfrak{S}_{l'a}^{(t')} + C_{a'b}^{-1} \mathfrak{S}_{l'b}^{(t')}),$$

the symmetry operation having also shifted t' to $C^{-1}t'$. Then summation over C and use of Eq. (16) gives the result above.

Appendix II

It is to be proved that the use of the smaller matrix

$$\mathfrak{G}^0 = \mathfrak{G}_{22} - \mathfrak{G}_{21}\mathfrak{G}_{11}^{-1}\mathfrak{G}_{12}$$

is equivalent to putting the force constants in the rows and columns denoted by 1 equal to infinity.

The kinetic energy is $2T = (d\mathfrak{R}'/dt)\mathfrak{G}^{-1}(d\mathfrak{R}/dt)$. Putting certain force constants equal to infinity is equivalent to making certain bonds rigid. They may therefore be dropped out of the kinetic energy. Let $\mathfrak{G}^{-1}$ be written in terms of submatrices as

$$\mathfrak{G}^{-1} = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}$$

in which 1 represents the rows and columns to be dropped. The kinetic energy which will be used is then $2T = (d\mathfrak{R}_2'/dt)X_{22}(d\mathfrak{R}_2/dt)$, with all the coordinates with infinite force constants dropped from $\mathfrak{R}$ to give $\mathfrak{R}_2$. If the original $\mathfrak{G}$ is written

$$\mathfrak{G} = \begin{pmatrix} \mathfrak{G}_{11} & \mathfrak{G}_{12} \\ \mathfrak{G}_{21} & \mathfrak{G}_{22} \end{pmatrix},$$

then since $\mathfrak{G}\mathfrak{G}^{-1} = E$,

$$\mathfrak{G}_{11}X_{12} + \mathfrak{G}_{12}X_{22} = 0,$$
$$\mathfrak{G}_{21}X_{12} + \mathfrak{G}_{22}X_{22} = E_{22},$$

or

$$X_{12} = -\mathfrak{G}_{11}^{-1}\mathfrak{G}_{12}X_{22}.$$

so that

$$X_{22} = (\mathfrak{G}_{22} - \mathfrak{G}_{21}\mathfrak{G}_{11}^{-1}\mathfrak{G}_{12})^{-1}.$$

The secular equation in the form $|\mathfrak{F}_{22} - X_{22}\lambda| = 0$ can then be multiplied through by $X^{-1}{}_{22} = \mathfrak{G}^0$ to give the desired result.

On the Theory of Antiferromagnetism*

J. H. Van Vleck
Harvard University, Cambridge, Massachusetts
(Received October 19, 1940)

Expressions for the susceptibility of a medium with antiferromagnetic exchange coupling are derived by means of the Dirac vector model, with an accuracy comparable to that in the usual first approximation of the Weiss-Heisenberg theory of ferromagnetism. The formulas yield a Curie temperature T_C above which the internal field disappears and at which the susceptibility χ is a maximum. The calculated value of χ is $\frac{2}{3}$ as large at $T=0$ as at $T=T_C$. This prediction accords nicely with experimental data on MnO, but not on MnSe. In an antiferromagnetic medium, the internal field should probably be treated as randomly directed, and our results may be regarded as generalizations and combinations of formulas previously obtained by Néel and Hulthén and by Bitter for the cases of inner fields respectively perpendicular and parallel to the applied field. A slight accentuation of the perpendicular alignment by very powerful applied fields explains the dependence of the susceptibility on field strength observed by Bizette, Squire, and Tsai in MnO when $T<T_C$.

Introduction.—An increasing interest is being manifested in so-called antiferromagnetic media, i.e., substances in which the exchange coupling favors antiparallel rather than parallel alignment of the spins. However, the formulas yielded in the first approximation by the usual Heisenberg model do not appear to have been explicitly given in this case of the opposite sign from usual, although they are contained implicitly in a combination of the work of Néel[1] and of Bitter.[2] It is the purpose of the present paper to derive them.

The Dirac formalism[3] shows that exchange interactions are equivalent to an interatomic potential

$$V_{ij}=-\tfrac{1}{2}J[1+4\mathbf{S}_i\cdot\mathbf{S}_j], \qquad (1)$$

where $\mathbf{S}_i$ and $\mathbf{S}_j$ are, respectively, the spin angular momentum vectors[4] of atoms i and j, measured in multiples of the quantum unit $h/2\pi$, and J is the exchange integral. For an antiferromagnetic medium, J is negative. We make the usual approximations of assuming the exchange coupling can be neglected except between atoms which are nearest neighbors, and that the electrons are effectively in S states, with a common exchange integral J.

The effective potential V_i to which a given atom i is subjected may be obtained by summing over the various atoms j which are its immediate neighbors. It has been shown by Stoner[5] that the formulas characteristic of the first approximation of the Heisenberg theory can be obtained by using mean values for the positions of the spins of the atoms contiguous to a given atom. Thus, apart from a trivial additive constant, we may take

$$V_i=-2J\mathbf{S}_i\cdot\textstyle\sum_j\bar{\mathbf{S}}_j=-2zJ\mathbf{S}_i\cdot\bar{\mathbf{S}}_j, \qquad (2)$$

where z is the number of neighbors possessed by a given atom. These neighboring atoms have the same alignment on the average, so that summation over j in (2) is equivalent to multiplication by z of the mean value for any one atom j.

Ordinarily, i.e., in the ferromagnetic case, the taking of mean values in the fashion (2) is equivalent to assuming that all states with the same resultant spin for the entire crystal have the same energy. However, in antiferromagnetic media, this approximation must be used in a somewhat different way, as follows, which allows no particular interpretation in terms of the total crystalline spin, but which nevertheless represents a degree of accuracy comparable with the usual first approximation in the ferromagnetic case.

Adaptation of the Dirac vector model to alter-

* Presented at the Symposium on the Structure of Molecules and Aggregates at the Fifth Annual Symposium of the Division of Physical and Inorganic Chemistry of the American Chemical Society, Columbia University, New York, December 30, 1940 to January 1, 1941.

[1] L. Néel, Ann. de physique **17**, 64 (1932); **5**, 256 (1936).

[2] F. Bitter, Phys. Rev. **54**, 79 (1937).

[3] P. A. M. Dirac, Proc. Roy. Soc. **123**, 714 (1929); for the application to ferromagnetism, see Chapter XII of my *Electric and Magnetic Susceptibilities*.

[4] Throughout the paper, all expressions printed in boldface type are vectors.

[5] E. C. Stoner, Proc. Leeds Phil. Soc. **2**, 56 (1930); Phil. Mag. **10**, 43 (1930).

nating sub-lattices.—We must divide the crystal into sub-lattices a and b, such that at the absolute zero all the spins of the atoms of one sub-lattice point in one direction, and those of the other point in the diametrically opposite direction. The sub-lattices a and b are to contain equal numbers of atoms, and the nearest neighbors of an atom on a are to belong entirely to b, and vice versa. Our analysis thus applies only to simple and body-centered cubic structures, and not to crystals such as those of the face-centered cubic or hexagonal types, wherein nearest neighbors of a given atom can also sometimes be nearest neighbors of each other. For these more involved cases, some sort of a more complicated procedure, which we shall not consider, would be involved,—perhaps dividing the lattice into more than two sub-lattices.

Henceforth we will let the subscripts i and j refer to atoms belonging to the two different sub-lattices. In the absence of an applied field $\mathbf{H}$, we may suppose the mean values of the spins $\mathbf{S}_i$ and $\mathbf{S}_j$ to have the opposite directions. The introduction of the field $\mathbf{H}$ will give a displacement $\delta\mathbf{S}_i$ or $\delta\mathbf{S}_j$ to the spin of a typical atom i or j. These displacements we can, without loss of generality, consider to be parallel and proportional to $\mathbf{H}$, since any possible distortions perpendicular to, or nonlinear in, $\mathbf{H}$ will not enter for our purposes, inasmuch as we will consider only the powder susceptibility, and neglect saturation effects. We can further take $\delta\mathbf{S}_i=\delta\mathbf{S}_j=\delta\mathbf{S}$, the justification being that for small fields the susceptibility of a given atom is the same regardless of whether the applied field is rotated by 180° relative to the internal field. We thus have

$$\bar{\mathbf{S}}_i=\mathbf{S}_0+\delta\mathbf{S}, \quad \bar{\mathbf{S}}_j=-\mathbf{S}_0+\delta\mathbf{S}. \tag{3}$$

Here the term $\mathbf{S}_0$ represents what may be termed "spontaneous antiferromagnetism," as it is oppositely directed for the two types of sub-lattices. The critical temperature above which $\mathbf{S}_0$ vanishes is the Curie temperature. The moment per unit volume is

$$\mathbf{M}=\tfrac{1}{2}gN\beta[\bar{\mathbf{S}}_i+\bar{\mathbf{S}}_j]=Ng\beta\delta\mathbf{S}, \tag{4}$$

where β is the Bohr magneton $he/4\pi mc$, N is the number of atoms per unit volume, and g is the usual Landé factor, for our purposes practically 2.

A given atom i or j also experiences, besides the potential (2), an energy $-g\beta\mathbf{H}\cdot\mathbf{S}_i$ or $-g\beta\mathbf{H}\cdot\mathbf{S}_j$ due to the applied field $\mathbf{H}$. Hence, using (2) and (3), we see that with our type of approximation, the exchange effects are equivalent to taking the total effective field acting on a given atom as

$$\begin{aligned}\mathbf{H}^{\pm}_{\text{eff}}&=\mathbf{H}+2Jzg^{-1}\beta^{-1}(\mp\mathbf{S}_0+\delta\mathbf{S})\\&=\mathbf{H}+2Jzg^{-1}\beta^{-1}(\mp\mathbf{S}_0)+(2Jz\mathbf{M}/Ng^2\beta^2).\end{aligned} \tag{5}$$

The plus or minus sign in the left side of (5) is to be used according as we are dealing with the field exerted on atom i or j. The opposite sign choice applies to the right rather than to the left side of (5) because the neighbors belong to the opposite category from the atom itself.

In accordance with standard magnetic theory[6] for the susceptibility of atoms subject to an effective field H_{eff}, we see that the magnetic moment $M=|\mathbf{M}|$ per unit volume in the direction of the applied field is given by the expression

$$\begin{aligned}M=\tfrac{1}{2}gSN\beta[&\cos(\mathbf{H}^{+}_{\text{eff}},\mathbf{H})B_S(y^+)\\&+\cos(\mathbf{H}^{-}_{\text{eff}},\mathbf{H})B_S(y^-)],\end{aligned} \tag{6}$$

where $B_S(y)$ is the Brillouin function

$$B_S(y)=\frac{2S+1}{2S}\coth\left(\frac{2Sy+y}{2S}\right)-\frac{1}{2S}\coth\left(\frac{y}{2S}\right), \tag{7}$$

and where we use the notation

$$y^{\pm}=|\mathbf{H}^{\pm}_{\text{eff}}|Sg\beta/kT. \tag{8}$$

Here S, to be carefully distinguished from the vector $\mathbf{S}$, is the spin quantum number for any atom of our ensemble. For the common case $S=\frac{1}{2}$, the function $B_S(y)$ reduces to $\tanh y$. In writing (6), we have utilized the fact that we must consider our crystal as an equal mixture of the two types i and j. When we apply the formulas separately to the two types, by isolating the corresponding contributions to (6), we have

$$\begin{aligned}|-\mathbf{S}_0+\delta\mathbf{S}|&=SB_S(y^-),\\|\mathbf{S}_0+\delta\mathbf{S}|&=SB_S(y^+).\end{aligned} \tag{9}$$

The two relations embodied in (9) are consistent

[6] See, for example, J. H. Van Vleck, *The Theory of Electric and Magnetic Susceptibilities* (Oxford, 1932), p. 257.

and equivalent, as the neglect terms beyond the first degree in the applied field $\mathbf{H}$. For $\mathbf{H}=0$, they both reduce to

$$|\mathbf{S}_0| = SB_S(y_0) \quad \text{with} \quad y_0 = 2zS|J\mathbf{S}_0|/kT. \quad (10)$$

Equations (10) determine the "antiferromagnetic spontaneous field" $\mathbf{S}_0$. Except for the fact that now $|J|=-J$, these equations are entirely analogous to the corresponding relations in the Weiss-Heisenberg theory for a ferromagnetic medium. Thus, for given $|J|$, the internal field has the same magnitude in the ferromagnetic and antiferromagnetic cases, but with, of course, the difference that in the latter, the field, instead of consistently preserving the same direction, is staggered as we go from atom to atom.

The Curie point.—The Curie temperature is that above which (10) ceases to have a real nonvanishing solution for $\mathbf{S}_0$. Analytically this is expressed[7] by the fact that $\mathbf{S}_0=0$ is a double root of (10) at $T=T_C$. For small values of the argument, we have

$$SB_S(y) = \tfrac{1}{3}(S+1)y. \quad (11)$$

Hence the formula for T_C is

$$T_C = \tfrac{2}{3}|J|zk^{-1}S(S+1). \quad (12)$$

This is the same expression for T_C as in the first approximation of the standard ferromagnetic theory,[7] except that an absolute value sign is now necessary since $J<0$. The spontaneous internal field present below the Curie point manifests itself in an antiferromagnetic medium not in saturation, enormous susceptibilities, etc., as in a true ferromagnetic, but instead, primarily, in an altered variation of the susceptibility with temperature, as we shall see later.

$T>T_C$.—Above the Curie point, we may take $\mathbf{S}_0=0$ in (5), as there is no spontaneous field. The cosines in (6) become unity. Also, we may use (11). Hence on substituting (11), (8), in (6), and solving for M, we have

$$M = Ng^2\beta^2S(S+1)H/3k(T+T_C), \quad (13)$$

with T_C as in (12). The corresponding formula for a ferromagnetic medium has $T-T_C$ rather than $T+T_C$ in the denominator. The sign reversal is due to the fact that in the present, or antiferromagnetic case, we have $|J|=-J$.

$T<T_C$, $\mathbf{S}_0\|\mathbf{H}$. Below the Curie point, there is the spontaneous antiferromagnetism. We cannot, however, assume that $\mathbf{S}_0$ is necessarily parallel or antiparallel to $\mathbf{H}$, as $\mathbf{S}_0$ may be oblique to $\mathbf{H}$. We shall first, however, consider the case that $\mathbf{H}$ and $\mathbf{S}_0$ are collinear, so that the cosines in (6) are ± 1. The alternation in sign expresses the staggering of the inner field, which is now supposed larger than the applied field, and which so determines the direction of $\mathbf{H}_{eff}$. We neglect saturation effects, and so may expand the Brillouin functions in (6) as Taylor's series in $H=|\mathbf{H}|$, keeping only the constant term. Hence using (5) we may write

$$B_S(y^{\pm}) = B_S(y_0) \pm \left(\frac{g\beta S}{kT}\right)\left(H+\frac{2JzM}{Ng^2\beta^2}\right)B_S'(y_0), \quad (14)$$

where $y^{\pm}$, y_0 are defined as in (8), (10) and where the prime denotes the derivative of the Brillouin function with respect to its argument. Substitution of (14) in (6) now yields

$$M = \frac{Ng^2\beta^2S^2B_S'(y_0)H}{k[T+3T_C(S+1)^{-1}SB_S'(y_0)]}(\mathbf{M}\|\mathbf{S}_0). \quad (15)$$

For the particular case $S=\frac{1}{2}$, Eq. (15) may be written

$$M = \frac{Ng^2\beta^2H}{4k[T_C+(1-4\mathbf{S}_0^2)^{-1}T]}, \quad (16)$$

while for the classical limit $S\to\infty$, $\beta\to 0$, it becomes, on setting $Sg\beta=\mu$,

$$M = \frac{N\mu^2H}{3kT_C+kT[1-(\mathbf{S}_0/S)^2-\frac{2}{3}(T/T_C)]^{-1}}. \quad (17)$$

The expression $4\mathbf{S}_0^2$ appearing in (16), or $(\mathbf{S}_0/S)^2$ in (17), has the interpretation of being the square of the ratio of the values of the spontaneous inner field at the temperature T and at the absolute zero.

$T<T_C$, $\mathbf{H}\perp\mathbf{S}_0$. If $\mathbf{H}$ is applied perpendicular to the spontaneous moment $\mathbf{S}_0$, the field acting on any given atom will consist of two parts, *viz.*, the original inner field $\mathbf{H}_x=\mp 2Jzg^{-1}\beta^{-1}\mathbf{S}_0$ directed along, say, the x axis, and the applied plus internal induced field $\mathbf{H}_z=\mathbf{H}+2Jzg^{-1}\beta^{-1}\delta\mathbf{S}$ di-

[7] Cf., for instance, reference 6, p. 332.

rected along the z axis. As $\mathbf{H}_z$ is small in magnitude compared to $\mathbf{H}_x$, we may take the cosine factors in (6) to both be $|\mathbf{H}_z|/|\mathbf{H}_x|$, and replace $y^{\pm}$ by y_0 in the arguments of the Brillouin function. Then the mean moment per unit volume is by (6) and (10)

$$M=\frac{N|\mathbf{H}_z|}{|\mathbf{H}_x|}g\beta|\mathbf{S}_0|=\frac{Ng^2\beta^2|\mathbf{H}+2Jzg^{-1}\beta^{-1}\delta\mathbf{S}|}{2|J|z}.$$

On using (4), eliminating J by means of (12), and solving for M, we have the simple result

$$M=Ng^2\beta^2S(S+1)H/6kT_C. \quad (\mathbf{M}\perp\mathbf{S}_0) \quad (18)$$

Below the Curie point, the moment thus becomes independent of temperature when the external field is applied perpendicular to the inner field.

Random orientation of the internal field.—Actually, we cannot use the idealization that the internal field is parallel or perpendicular to the applied one. However, the correct result will be obtained if we assume that it is perpendicular and parallel, respectively, two-thirds and one-third of the time. This procedure is not a rough approximation, but instead is accurate as long as we are neglecting saturation effects. Namely, we can resolve $\mathbf{H}$ into components parallel and perpendicular to $\mathbf{S}_0$. The corresponding contributions to the moment are easily seen to be additive, and give $M_{||}\cos^2(\mathbf{H},\mathbf{S}_0)$ and $M\sin^2(\mathbf{H},\mathbf{S}_0)$, respectively. The direction of the internal field may obviously be taken to be random in a powder. Also, even in a single crystal we can replace $\cos^2(\mathbf{H},\mathbf{S}_0)$ and $\sin^2(\mathbf{H},\mathbf{S}_0)$ by mean values $\frac{1}{3}$ and $\frac{2}{3}$, respectively, inasmuch as there are always three equivalent directions in a cubic crystal. So we can safely write for the susceptibility

$$M=\tfrac{1}{3}M_{||}+\tfrac{2}{3}M_{\perp}, \quad (19)$$

where $M_{||}$ and $M_{\perp}$ are given by (15) and (18), respectively, when $T<T_C$. At the critcial temperature, (19) joins continuously with the moment given by (13) for $T>T_C$, as should be the case. Namely, near T_C, we can compute $B'(y_0)$ by means of (11), so that (15) becomes identical with (13) at $T=T_C$, while (18) obviously does also. A little inspection[8] shows that $M_{||}$ vanishes at $T=0$, and increases monotonely as T is raised from 0 to T_C. On the other hand, $M_{\perp}$ is constant below $T=T_C$. As $dM/dT<0$ for $T>T_C$, we see that the susceptibility has a maximum at $T=T_C$, and is two-thirds as large at the absolute zero as at the Curie temperature. The situation is illustrated in Fig. 1 for the limiting cases $S=\frac{1}{2}$, and $S=\infty$. The fact that the curves do not differ very much for these two values of S shows that the results, when expressed in "reduced units" such as those we use in Fig. 1 are not sensitive to the size of the atomic spin. The exact shape of the curves should not be taken too literally, as it is well known that the "first approximation" or method of mean values which we have used is not really accurate, especially at low temperatures. However, the results $M_{\perp}$=constant for $T<T_C$ and $M_{||}=0$ at $T=0$ are valid as long as one uses any method based on the molecular field. For instance, one could use the molecular fields involved in the Bethe-Peierls[9] theory, which are nonlinear in M and so more general than our procedure. Hence the trend of our formulas is probably more than qualitatively correct for the model in question.

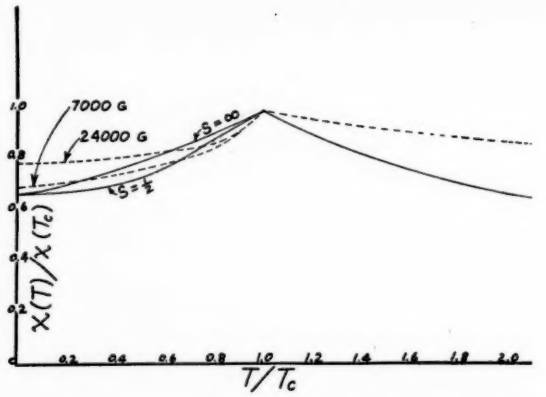


FIG. 1. The susceptibility as a function of the reduced temperature T/T_C. The solid curves are computed from the theoretical formulas (15), (18), (19), while the dashed lines give the experimental results of Bizette, Squire, and Tsai on MnO.

[8] The quantity $B_S'(y_0)$ in (15) vanishes more rapidly at $T=0$ than does T itself. Correspondingly the denominators of (16) and (17) involve the reciprocals of expressions vanishing more rapidly than T. These statements apply regardless of whether one uses our molecular field model or the Bloch spin waves. In the case $S=\frac{1}{2}$, for example, the factor $1-4S_0^2$ appearing in (16) vanishes exponentially in the former, and as $T^{\frac{3}{2}}$ in the latter (cf. reference 6, p. 338).

[9] R. Peierls, Proc. Camb. Phil. Soc. **32**, 477 (1936); P. Weiss, Phys. Rev. **55**, 673 (1939).

Comparison with experiment.—The predictions of formula (19) taken in conjunction with (15) and (18), are in fairly good agreement with experiment. The most notable characteristic of an antiferromagnetic medium is a certain critical or "Curie" temperature at which the susceptibility reaches a maximum, and precisely this behavior is found in certain substances, such as MnO. Bizette, Squire, and Tsai[10] find that the susceptibility of MnO is 0.7 as large at $T=14°$ as at the Curie temperature. The value at 0° probably is slightly lower than that at 14°K, so that the ratio of the extrapolated susceptibility at $T=0$ to that at the Curie point has, in weak fields, almost exactly the theoretical value 0.67. Above the Curie point they find that the susceptibility is inversely proportional to $T+\Delta$ where the constant Δ is about five times T_C instead of being equal to T_C as required by (13). Their results are indicated by the dashed lines in Fig. 1. In the case of MnSe, Squire[10] finds that the susceptibility at liquid-air temperatures is less than half that at the Curie point, so that the total drop is greater than the 33 percent yielded by the theory. Antiferromagnetic behavior is also reported by Foex and Graff[11] for Cr_2O_3 and CrSb. The susceptibilities of these two substances at low temperatures are, respectively, about 85 and 30 percent of the maximum values characteristic of the Curie point.

Exact agreement with our theory cannot be expected, as our method based on the taking of mean values represents a degree of accuracy comparable with the first approximation of Heisenberg's theory, and cannot be regarded as quantitatively correct. Also, especially, there are known[12] to be other causes of feeble paramagnetism besides exchange demagnetization, so that the theory may not embody all the ingredients actually present. In particular, if part of the susceptibility of MnO is due to other causes, and if this other portion is more or less independent of temperature, an explanation may be obtained of why the susceptibility falls off less rapidly above the Curie point than predicted by (19), but then it is hard to understand why the magnetization falls off so much at low temperatures, since, with the superposition of the other contributions, the susceptibility should diminish by less than 33 percent in going from $T=T_C$ to $T=0$.

Dependence on field strength.—According to our theory, the susceptibility should vary somewhat with field strength, as the susceptibility is higher when the inner field is perpendicular rather than parallel to the applied field. Hence, as already suggested by Néel,[1] a sufficiently powerful external field should be adequate to rotate the internal fields of the elementary domains so that they are perpendicular to the applied field instead of being random. In a very large field, the susceptibility should remain constant below T_C since (18) gives a constant M for $T<T_C$. The coercive fields necessary to rotate the elementary domains should be much greater than in a true ferromagnetic, for the decrease in energy of the aligned over the random orientation is now $\frac{1}{3}(M_\perp - M_{||})H$ instead of $\frac{1}{3}g\beta N|\mathbf{S}_0|H$, and is, therefore, proportional to H^2 rather than H. As $M_\perp - M_{||}$ differs from $g\beta NS_0$ by a factor of the order H/H_{int}, and a true ferromagnetic medium has an internal field $H_{\text{int}}\sim 10^5$ gauss, the necessary field should be of the order 10^5 gauss rather than a matter of a few gauss as in a ferromagnetic material. Figure 1 shows that Bizette, Squire, and Tsai[10] do indeed observe a different susceptibility below the Curie point for fields of 7000 and 24,000 gauss. A particularly striking fact is that the difference between the susceptibilities for the two fields shrinks to zero at the Curie point. Precisely this behavior is predicted by the formulas, for (15) and (18) become identical at $T=T_C$, and do not differ materially unless T is appreciably less than T_C, the maximum difference, of course, being at $T=0$.

It may be noted that in MnS and MnSe, Squire[10] finds hysteresis and dependence on field strength even in fields of a few gauss, with the phenomenon persisting even above the Curie point. Effects of this type presumably have nothing to do with those in large fields described in the preceding paragraph, and just how they are to be explained is not at present clear. Dr. Squire suggests to the writer that possibly supercooling is the cause, as measurements on the specific heat of MnSe show anomalies resulting from supercooling.

[10] H. Bizette, C. F. Squire and Tsai, Comptes rendus **207**, 449 (1938); C. F. Squire, Phys. Rev. **56**, 922 (1939).
[11] M. G. Foex and M. Graff, Comptes rendus, **209**, 106 (1939).
[12] Cf. reference 6, p. 347.

Relation to other work.—In closing, we must by all means mention that many of our formulas have already been published in one form or another. Néel,[1] in particular, stressed the need of considering the possibility of the inner field being perpendicular to the applied one, and gave (18) for the classical limit $S=\infty$. Also, he pointed out that in the parallel case the susceptibility vanishes at $T=0$, at least in classical theory, and noted that (13) applies for $T\gg T_C$. Néel, however, does not obtain a linear relation between $1/\chi$ and T unless T is well above T_C, and no doubt (13) is an excessive idealization in this respect. The expression (16) for the parallel alignment in the particular case $S=\frac{1}{2}$ has already been deduced by Bitter in a somewhat phenomenological fashion.[2] A properly combined and general quantal derivation for the two alignments at an arbitrary temperature does not appear to have previously been given.

An entirely different, and probably more rigorous, mathematical approach to the calculation of antiferromagnetism has been developed by Hulthén.[13] He used a semiclassical analog[14] of the Bloch method of spin waves, so that his calculations apply only to very low temperatures. Also he considered only the case that $\mathbf{S}_0$ is perpendicular to $\mathbf{H}$. The limiting value $\chi_{\perp 0}$ of the susceptibility $\chi_{\perp}=M_{\perp}/H$ which he obtains at $T=0$ agrees with (18). He derives the result that $\chi_{\perp}$ should approach this limit in the fashion $\chi_{\perp}/\chi_{\perp 0}=(1-aT^2)$. The corresponding formula of Néel has the structure $\chi_{\perp}/\chi_{\perp 0}=1-bT^{-1}(e^{b/T}-1)^{-1}$. The difference between the expressions for $\chi_{\perp}$ which Néel obtains at $T=0$ and $T=T_C$ is negligible ($1\frac{1}{2}$ percent) so that the amount of error introduced by using (18) is trivial. According to Hulthén, however, there is rather more dependence of $\chi_{\perp}$ on T than predicted by Néel or by (18), which gives none. This discrepancy no doubt arises because Hulthén's method includes fluctuation effects not contemplated in the molecular field model. It probably does not prevent (18) from being a reasonably good approximation, for Hulthén's results apply only at very low T, where aT^2 is small compared with unity in his factor $(1-aT^2)$. In making this statement regarding (18), we are guided partly by analogy with the corresponding situation in ferromagnetic media, where it is well known that the formulas derived by the molecular field model give the wrong approach of the saturation magnetization to its limiting value at $T=0$, but still represent remarkably well the variation with temperature over the interval $T=0$ to $T=T_C$ considered as a whole.[15]

In the discussion of antiferromagnetism in the literature, reference is occasionally made to a paper by Landau.[16] His treatment, however, is based on a highly phenomenological model in which the coupling is arbitrarily supposed ferromagnetic in certain directions and antiferromagnetic in others. Hence no simple connection is to be expected between his formulas and those which we have given or quoted, all of which are based on the simple, isotropic exchange potential (1).

To our degree of approximation, the results for the case that $\mathbf{S}_0$ and H are parallel would be the same if one used the Ising[17] rather than the true vector model. Peierls[9] has shown that there is a mathematical similarity between ferromagnetic calculations made by using the Ising model with $S=\frac{1}{2}$, and the well-known order-disorder theory of alloys.[18] This formal resemblance to the latter persists even with antiferromagnetic coupling if one is careful to introduce a staggered internal field, which is a new physical feature. The possibility of $\mathbf{S}_0$ being aligned obliquely to H, however, introduces a directional element which has no simple counterpart in the Ising or alloy theory.

[13] L. Hulthén, Proc. Amst. Acad. Sci. **39**, 190 (1936).

[14] Due to G. Heller and H. A. Kramers, Proc. Amst. Acad. Sci. **37**, 378 (1934).

[15] Cf. reference 6, p. 338.

[16] L. Landau, Phys. Zeits. Sowjetunion **4**, 675 (1934).

[17] E. Ising, Zeits. f. Physik **31**, 253 (1925). The Ising model has $S_{zi}S_{zj}$ rather than $S_i\cdot S_j$ in (1). However, only S_{xi} and S_{xj} contribute to (2) if both $\mathbf{H}$ and $\mathbf{S}_0$ are parallel to the z axis.

[18] N. F. Mott and H. Jones, *Properties of Metals and Alloys*, p. 28 ff. Peierls (reference 9) discussed primarily the analogy to the theory of adsorption, but there is also a corresponding parallelism to alloy theory.

Overvoltage and the Structure of the Electrical Double Layer at a Hydrogen Electrode*

GEORGE E. KIMBALL, SAMUEL GLASSTONE, AND ALVIN GLASSNER
Columbia University, New York, New York, and Princeton University, Princeton, New Jersey

(Received October 9, 1940)

The theory of absolute reaction rates applied to electrode processes leads to the equation

$$s=(kT/h)\exp(-\Delta F^*/RT)\exp(\alpha V\mathfrak{F}/RT)$$

for the specific rate of discharge of ions. The question arises as to whether V should be the actual potential at the electrode or only the overvoltage. In the case of hydrogen ion discharge, it is found experimentally that the overvoltage should be used. On the other hand, the theory in its simplest form requires V to be the total potential. This paradox can be resolved by postulating the existence of two different electrical double layers at the electrode surface, and two corresponding energy barriers over which the protons must pass. Provided that the barrier nearer the electrode is the higher, the overvoltage is essentially established across this layer, while the variation in equilibrium potential caused by variations in the hydrogen ion concentration of the solution is established across the outer double layer. Since the rate of discharge is determined by the potential difference across the inner double layer, the rate is determined by the overvoltage, and not by the total potential. The nature of these two energy barriers is also discussed. Both barriers may correspond to transfer of protons from water molecule to water molecule, or one may correspond to the actual discharge process in which a neutral hydrogen atom is formed. It does not yet seem to be possible to decide between these two alternatives, either theoretically or experimentally, but it seems most probable that both barriers correspond to proton transfers from water molecule to water molecule.

IN applying the theory of absolute reaction rates to electrode processes,[1,2] the equation giving the specific rate of the discharge of ions, or of the reverse process, takes the form

$$s=(kT/h)\exp(-\Delta F^*/RT)\exp(\alpha V\mathfrak{F}/RT)$$

and the question arises as to whether V should be the actual potential at the electrode or only the overvoltage; i.e., the difference between the actual potential and the reversible potential in the same solution.[3] For the treatment of hydrogen ion discharge[1] it is necessary that V should represent the overvoltage, for otherwise the equation would lead to the conclusion that the hydrogen overvoltage varies with the hydrogen ion concentration, whereas the experimental evidence indicates that the former is independent of the latter, at least at low current densities. On the other hand, it must be admitted that the whole, or almost the whole, potential difference between a metal and a solution operates between the metal itself and an adsorbed ion layer,[2] and hence it is to be expected that the rate of passage of an ion from this layer to the electrode should be affected by a fraction α of the whole potential difference, and not merely of the overvoltage.

To resolve this apparent paradox it is necessary to postulate that the overvoltage is set up at a different double layer from that responsible for the reversible potential,[4] or at least for that part of the reversible potential which varies with the *p*H of the solution. If the entire potential difference at the electrode is caused by a single layer of adsorbed ions and its electrical image in the electrode, this is not possible. There must, therefore, be at least two layers of ions next to the electrode which are not electrically neutral. We shall denote the outer layer, farther from the electrode, as the "solution layer." The inner layer we shall call the "adsorbed layer," but this does not necessarily imply that this layer is in direct contact with the electrode surface.

As a proton passes from the solution to the

* Presented at the Symposium on the Structure of Molecules and Aggregates of Molecules at the Fifth Annual Symposium of the Division of Physical and Inorganic Chemistry of the American Chemical Society, Columbia University, New York, December 30, 1940 to January 1, 1941. Publication assisted by the Ernest Kempton Adams Fund for Physical Research of Columbia University.

[1] H. Eyring, S. Glasstone and K. J. Laidler, J. Chem. Phys. **7**, 1053 (1939).

[2] G. E. Kimball, J. Chem. Phys. **8**, 199 (1940).

[3] This point has also been raised by A. Frumkin, Acta Physicochimica, U.R.S.S. **12**, 481 (1940).

[4] Cf. reference 1, p. 1058, footnote.

electrode it must first reach the solution layer. From this it passes over an energy barrier to the adsorbed layer. Thence it must pass over another energy barrier and either be discharged or reach a new equilibrium position from which it can be discharged so easily that the discharge process may be regarded as instantaneous. In order to keep the argument general we shall refer to the equilibrium position reached by the proton after it leaves the adsorbed layer as the "electrode layer." The proton in the electrode layer may be in the form of an adsorbed hydrogen atom, or it may be part of an adsorbed water molecule. In either case we shall assume that the difference in potential between the electrode layer and the electrode itself is negligibly small. We shall also assume that the potential of the solution layer is the same as that of the solution proper.

The potential energy surface for a proton traversing these layers is shown in Fig. 1. The energy shown is that due to nonelectrical forces, and to it must be added the electrical energy $V\mathfrak{F}$ per mole, where V is the electrical potential and $\mathfrak{F}$ is the faraday. We shall denote by V_1 the potential difference between the adsorbed layer and the solution layer, and by V_2 the potential difference between the electrode layer and the adsorbed layer.

Let us denote the electrical charge per unit area in the solution layer by q_1; that in the adsorbed layer by q_2; and that in the electrode layer by q_3. Since we have assumed that the whole potential drop at the electrode occurs between the solution layer and the electrode layer, the sum $q_1+q_2+q_3$ must vanish. The system of three planes of charge may be regarded as two condensers in series. Let C_1 be the capacity of the double layer bounded by the adsorbed layer and the solution layer, and C_2 that of the double layer bounded by the adsorbed layer and the electrode layer. We may then express V_1 and V_2 in terms of q_1, q_2, and q_3 by the equations

$$V_1=-q_1/C_1 \tag{1}$$

$$V_2=q_3/C_2. \tag{2}$$

We may also express q_2 as

$$q_2=C_1V_1-C_2V_2. \tag{3}$$

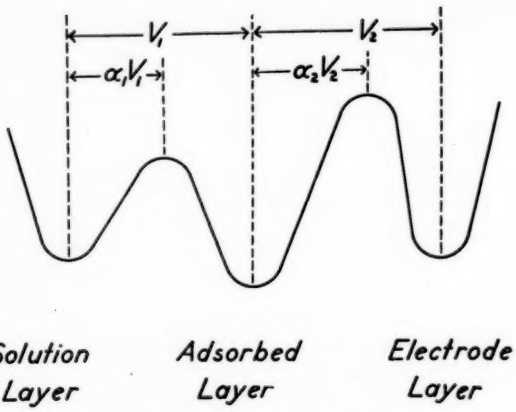


FIG. 1. Potential energy surface for protons approaching the electrode surface, showing the two barriers which must be surmounted.

Now all these layers must lie very close to the electrode surface, in fact within one or two diameters of a water molecule. In them we may have hydrogen (oxonium) ions and hydroxyl ions, but other ions can only enter at the expense of having lost some of their hydration energy. We shall, therefore, assume that the charge of the adsorbed ion layer is due entirely to hydrogen and hydroxyl ions. In the systems we shall consider (pH <7) this amounts to saying that the charge is due entirely to hydrogen ions.

We shall, therefore, assume that q_2 is proportional to the concentration of hydrogen ions in the absorbed ion layer. To a first approximation a_2, the hydrogen ion activity in this layer, is also proportional to the concentration, so that we may put a_2 proportional to q_2. Hence from Eq. (3) we may write

$$a_2=A_1V_1-A_2V_2 \tag{4}$$

where A_1 and A_2 are positive constants. The ratio A_1/A_2 is equal to C_1/C_2 or d_2/d_1, where d_1 is the distance from the solution layer to the adsorbed layer, and d_2 is the distance from the adsorbed layer to the electrode layer.

THE DISTRIBUTION OF POTENTIAL AT EQUILIBRIUM

Let us now consider the electrode at equilibrium. If a_1 and a_3 are the activities of the hydrogen ions in the solution layer and electrode layer, respectively, then the conditions of equi-

librium across the two barriers are

$$V_1=(RT/\mathfrak{F})\ln(a_1/a_2) \quad (5)$$

and

$$V_2=(RT/\mathfrak{F})\ln(a_2/a_3), \quad (6)$$

provided that the standard states are so chosen that the standard potentials are both 0. Since a_1 and a_3 are essentially the activities of the hydrogen ions in the solution and the hydrogen gas, respectively, they may be regarded as known. Equations (4), (5) and (6) are then sufficient to determine a_2, V_1, and V_2.

While these equations cannot be solved exactly, it is easy to obtain approximate solutions. Substitution of Eqs. (5) and (6) in Eq. (4) gives for a_2

$$a_2+(A_1+A_2)(RT/\mathfrak{F})\ln a_2 = A_1(RT/\mathfrak{F})\ln a_1+A_2(RT/\mathfrak{F})\ln a_3. \quad (7)$$

The substitutions $a_2=(A_1+A_2)(RT/\mathfrak{F})x$ and

$$y=(A_1/(A_1+A_2))\ln a_1 +(A_2/(A_1+A_2))\ln a_3-\ln(A_1+A_2)RT/\mathfrak{F}$$

reduce Eq. (7) to

$$x+\ln x=y. \quad (8)$$

Let the solution of this transcendental equation be

$$\ln x=\phi(y). \quad (9)$$

It is easily seen that to a first approximation for negative values of y

$$\phi(y)=y, \quad (y<0) \quad (10)$$

while for positive values of y

$$\phi(y)=(1-y^{-1})\ln y. \quad (y>0) \quad (11)$$

For $y>5$, this is practically a constant and we will write

$$\phi(y)=C \quad (y>5) \quad (12)$$

where C is almost constant.

Using these results we find, for $y<0$

$$\ln a_2=(A_1/(A_1+A_2))\ln a_1 +(A_2/(A_1+A_2))\ln a_3, \quad (13)$$

$$V_1=(A_2/(A_1+A_2))(RT/\mathfrak{F})\ln(a_1/a_3), \quad (14)$$

$$V_2=(A_1/(A_1+A_2))(RT/\mathfrak{F})\ln(a_1/a_3). \quad (15)$$

On the other hand, for $y>5$

$$\ln a_2=\ln(A_1+A_2)(RT/\mathfrak{F})+C, \quad (16)$$

$$V_1=(RT/\mathfrak{F})\ln a_1 -\ln(A_1+A_2)(RT/\mathfrak{F})-C, \quad (17)$$

$$V_2=(RT/\mathfrak{F})\ln(A_1+A_2)(RT/\mathfrak{F}) +C-\ln a_3. \quad (18)$$

The physical significance of these results is simple. If we imagine a very dilute acid solution in contact with an electrode covered with hydrogen at a very low pressure, then the concentration of hydrogen ions in the adsorbed layer will also be very small, and likewise q_2. The electrode potential is then established between a double layer consisting of the solution layer and the electrode layer. The values of V_1 and V_2 are then simply the total potential V times the ratio of the corresponding distance to the total distance between the solution layer and the electrode layer. This is the result expressed by Eqs. (14) and (15).

If the hydrogen ion concentration of the solution and the hydrogen gas pressure are now simultaneously increased in such a way that V remains constant, this result remains the same until eventually the concentration of hydrogen ions in the adsorbed layer becomes great enough to produce a charge in this layer large enough to affect V_1 and V_2. The accumulation of positive charge in this layer tends to resist the further increase of hydrogen ion concentration in this layer, and the value of a_2 becomes practically constant. When this occurs, further changes in the hydrogen ion concentration of the solution will affect only V_1, while changes in the gas pressure will affect only V_2.

The Distribution of Potential when Current is Flowing—Adsorption Control

If a current is flowing toward the electrode, the potential distribution depends on whether the slow step of the discharge process is the passage of protons from the solution layer to the adsorbed layer, or from the adsorbed layer to the electrode layer. In the former case we shall speak of adsorption control; in the latter, of discharge control.

If the electrode reaction is adsorption con-

trolled, the adsorbed layer is in equilibrium with the electrode layer, so that Eq. (6) is still valid. Instead of Eq. (5), however, we have an equation of the form[1,2]

$$I=a_0P\exp(-\alpha_1V_1\mathfrak{F}/RT), \qquad (19)$$

where I is the current density, α_1V_1 is the potential acting between the solution layer and the top of the energy barrier separating it from the adsorbed layer (cf. Fig. 1), P is a constant, and a_0 is the activity of the ionic or molecular species from which the protons principally come. If the current is carried chiefly by protons from oxonium ions, then a_0 is the same as a_1; while if the protons come chiefly from water molecules, a_0 is a constant.

If Eq. (19) is solved for V_1 and the result, together with Eq. (6) substituted in Eq. (4), we find

$$a_2+A_2(RT/\mathfrak{F})\ln a_2=A_2(RT/\mathfrak{F})\ln a_3 -A_1(RT/\alpha_1\mathfrak{F})\ln(I/Pa_0). \qquad (20)$$

The substitutions $a_2=A_2(RT/\mathfrak{F})x$ and

$$y=\ln a_3-(A_1/\alpha_1A_2)\ln(I/Pa_0)-\ln(A_2RT/\mathfrak{F})$$

reduce Eq. (20) to Eq. (8) so that again $\ln x=\phi(y)$, where $\phi(y)$ is the same function as before. Hence we have

$$\ln a_2=\ln(A_2RT/\mathfrak{F})+\phi(y)$$

and hence, from Eq. (6)

$$V_2=(RT/\mathfrak{F})\ln(A_2RT/\mathfrak{F}) -(RT/\mathfrak{F})\ln a_3+(RT/\mathfrak{F})\phi(y).$$

Combining this with Eq. (19) we find

$$\begin{aligned}V&=V_1+V_2\\&=(RT/\mathfrak{F})\ln(A_2RT/\mathfrak{F})-(RT/\mathfrak{F})\ln a_3\\&\quad+(RT/\mathfrak{F})\phi(y)-(RT/\alpha_1\mathfrak{F})\ln(I/Pa_0).\end{aligned}$$

The overvoltage, η, is given by

$$\begin{aligned}\eta&=(RT/\mathfrak{F})\ln(a_1/a_3)-V\\&=-(RT/\mathfrak{F})\ln(A_2RT/\mathfrak{F})+(RT/\mathfrak{F})\ln a_1\\&\quad-(RT/\mathfrak{F})\phi(y)+(RT/\alpha_1\mathfrak{F})\ln(I/Pa_0).\end{aligned}$$

Now if $y<0$, so that approximately $\phi(y)=y$

$$\eta=(RT/\mathfrak{F})\ln a_1-(RT/\mathfrak{F})\ln a_3 +(RT/\alpha_1\mathfrak{F})(1+(A_1/A_2))\ln(I/Pa_0). \qquad (21)$$

If $y>5$, so that approximately $\phi(y)=C$

$$\eta=(RT/\mathfrak{F})\ln a_1-(RT/\mathfrak{F})\ln(A_2RT/\mathfrak{F}) -(RT/\mathfrak{F})C+(RT/\alpha_1\mathfrak{F})\ln(I/Pa_0). \qquad (22)$$

Now it is expected that the value of α_1 is about $\frac{1}{2}$. Equation (22) is, therefore, in accord with experiment as far as the dependence of η on I is concerned. It predicts a dependence of η on a_1, however, whether a_0 is taken to be equal to a_1, or to be constant. Equation (21) is also in disagreement with the experimental evidence that η is independent of a_1. We conclude, therefore, that the rate-controlling step for the hydrogen electrode is *not* the passage of protons from the solution layer to the adsorbed layer.

The Distribution of Potential when Current is Flowing—Discharge Control

If the slow step in the discharge of hydrogen ions is the passage of protons from the adsorbed layer to the electrode layer, then the solution layer and the adsorbed layer are in equilibrium, and Eq. (5) is valid. Instead of Eq. (6), however, we now have an equation of the form

$$I=Pa_0\exp(-\alpha_2V_2\mathfrak{F}/RT), \qquad (23)$$

where α_2 has a similar significance to α_1, P is again a constant, and a_0 is now the activity of the molecular species in the adsorbed layer which is the chief source of protons.

Proceeding as before, we solve Eqs. (5), (23), and (4) by the aid of the substitutions $a_2=A_1(RT/\mathfrak{F})x$ and

$$y=\ln a_1+(A_2/\alpha_2A_1)\ln(I/Pa_0)-\ln(A_1RT/\mathfrak{F}).$$

We find

$$\eta=(RT/\mathfrak{F})\ln(A_1RT/\mathfrak{F})+(RT/\mathfrak{F})\phi(y) +(RT/\alpha_2\mathfrak{F})\ln(I/Pa_0).$$

For $y<0$ this takes the approximate form

$$\eta=(RT/\mathfrak{F})\ln a_1+(RT/\alpha_2\mathfrak{F}) \times(1+(A_2/A_1))\ln(I/Pa_0), \qquad (24)$$

while for $y>5$ we find

$$\eta=(RT/\mathfrak{F})\ln(A_1RT/\mathfrak{F})+(RT/\mathfrak{F})C +(RT/\alpha_2\mathfrak{F})\ln(I/Pa_0). \qquad (25)$$

Equation (24) predicts a dependence of η on a_1,

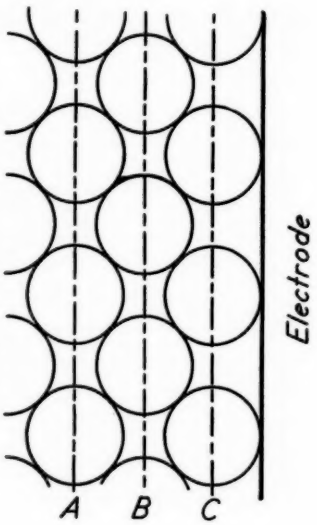


FIG. 2. Possible positions of the three charged layers. The solution layer is probably at *A*, the adsorbed layer at *B*, and the electrode layer at *C*; but possibly the solution layer is at *B*, the adsorbed layer at *C*, and the electrode layer at the electrode surface. The circles represent water molecules.

regardless of the interpretation of a_0, and is, therefore, not in accord with experiment. Equation (25), on the other hand, shows no dependence of η on a_1, provided that we take a_0 to be the activity of the water molecules in the adsorbed layer, as Eyring, Glasstone, and Laidler[1] have already suggested. Since Eq. (25) also gives the correct dependence of η on I, we conclude that under ordinary circumstances the discharge of hydrogen ions is controlled by the passage of protons from the water molecules of the adsorbed layer to the electrode layer.

We have, therefore, shown that the Eyring-Glasstone-Laidler mechanism of hydrogen overvoltage is essentially correct. Under the conditions for which Eq. (25) holds, the reversible potential is essentially established between the solution layer and the adsorbed layer, while the overvoltage is established between the adsorbed layer and the electrode layer.

It should be noted, however, that Eq. (25) can only hold when y is positive. For very low values of the hydrogen ion concentration, y will become negative, and Eq. (24) should be expected to hold. There is some experimental evidence[5] that Eq. (24) holds under some conditions, but the evidence is not clear, and will not be discussed here. It should be noticed, however, that Eq. (24) should hold at low current densities and low acidities. At low current densities, however, the back reaction, which has been neglected here, becomes important, so that experimentally this case would be difficult to observe. The most likely place to seek the transition would be in very weakly acid solutions.

[5] F. P. Bowden, Trans. Faraday Soc. **24**, 473 (1928); Proc. Roy. Soc **A126**, 107 (1929).

THE NATURE OF THE ENERGY BARRIERS

The foregoing arguments are independent of any assumption as to the location of the solution layer, the adsorbed layer, and the electrode layer, except that they must be sufficiently close to the electrode surface to prevent the entrance of ions other than hydrogen and hydroxyl ions into the adsorbed layer. Because of this condition, however, we should expect that these layers coincide with the layers of water molecules lying next to the electrode surface.

These layers are shown schematically in Fig. 2. The surface *C* contains the centers of the first layer of water molecules on the surface; the surfaces *B* and *A* the centers of the second and third layers. There are two possible locations of the layers considered in this paper. The three layers may be located along the planes *A*, *B*, and *C*; or the solution layer may be at *B*, the adsorbed layer at *C*, and the electrode layer on the electrode itself.

The latter possibility has the advantage that the adsorbed layer lies closest to the electrode, and therefore offers most resistance to the entrance of foreign ions. It also offers a simple explanation of why the activation energy for the passage of a proton from the adsorbed layer to the electrode layer is higher than that for the passage of a proton from the solution layer to the adsorbed layer. The latter is a simple transfer of a proton from one molecule to another, and should be relatively easy. The former, on the other hand, requires an electron jump from the electrode to the proton. This process would be expected to be more difficult, and hence have a higher activation energy.

The possibility that the three layers occupy the positions *A*, *B*, *C*, however, seems to have greater advantages. It is true that it is first necessary to postulate that the actual discharge process of a proton in layer *C* is very fast, but this is not impossible. On the other hand the calculations of Eyring, Glasstone, and Laidler[1]

of the entropy of activation for a proton transfer from one molecule of water to another agree with the observed entropy of activation for the discharge process, pointing to the fact that the rate controlling step is such a process. A more cogent reason for placing the three layers at A, B, and C is the similarity of the oxygen discharge process to the hydrogen discharge process. The explanation of this similarity, given by Eyring, Glasstone, and Laidler, can be carried over complete if the barriers needed for the present theory are both for proton transfers from water to water, but not if the rate determining step is the discharge of a proton.

The question of the exact location of these layers cannot be said to be completely settled, but it seems most probable that the location in the planes A, B, and C is correct.

JANUARY, 1941 JOURNAL OF CHEMICAL PHYSICS VOLUME 9

Optical Sensitizing of Silver Halides by Dyes

III. The Relation of Sensitizing to the Absorption Spectra and Constitution of Dyes*

S. E. Sheppard, R. H. Lambert and R. D. Walker
Eastman Kodak Company, Rochester, New York

(Received July 8, 1940)

Optical sensitizing is related to dye structure in general, not confined to particular classes of dyes. While sufficiently strong adsorption to silver halide, and relative insolubility of the adsorbed dye are essential, they are not sufficient conditions. The absorption spectra of adsorbed dyes correspond to their spectra in solution, and are affected in the same way by structural changes, but are not identical therewith. The displacement of the spectrum should be referred to the absorption of the dye in the gaseous state at low pressures; examples are given. This displacement corresponds to an adsorption energy available for sensitizing, but this energy difference may not be available, and is not generally sufficient. For "molecular" sensitizing, planarity of the dye molecule appears to be essential, and the following conditions are deduced as necessary and probably sufficient for sensitizing: (i) Planar configuration of dye molecules, respectively of adsorbed dye ions, (ii) edge-on adsorption of planar molecules oriented possibly orthogonally but more probably at an angle of 70° to a (111) plane of the crystal, and (iii) electronic transition in dye ion or dipole on absorption of a photon polarized in an azimuth defined by (i) and (ii). On these bases a hypothesis of coplanar coupling of electronic displacements in the dye and in a congruent plane of the silver halide lattice is advanced. It is suggested that this is a key factor in optical sensitizing by dyes, both for molecular and aggregate sensitizing. Certain other relations of constitution to sensitizing are discussed, including the influence of nuclear (perichrome) changes and of substitution.

THE fact that the sensitizing spectrum of dyes is related to their absorption spectrum —usually measured in alcohol or water—and again that both these are related to the molecular constitution, was realized early. Extensive investigations by H. W. Vogel, by V. Schumann, and especially by J. M. Eder and E. Valenta of Vienna and their collaborators[1] established such general relations, but in an equivocal and uncertain form which has not been greatly changed by the modern period in optical sensitizing, dating from the introduction of the cyanine or polymethine dyes.[2]

From the earlier investigations it can be concluded that practically all classes of dyes sufficiently soluble in aqueous solutions can in particular examples effect some degree of optical sensitizing. This is a sufficiently important fact to be emphasized, *viz.*, that *optical sensitizing per se is affiliated to dye structure in general*, rather than to special classes of dyes—with reservations in regard to dyes insoluble in water or alcohol.†

* Communication No. 774 from the Kodak Research Laboratories.

[1] J. M. Eder and E. Valenta, *Beiträge zur Photochemie u. Spektralanalyse* (W. Knapp, Halle, Vienna, 1904).

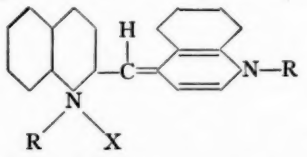
[2] A. Miethe and I. Traube, cf., C. E. K. Mees, J. Frank. Inst. **201**, 525 (1926).

† Dyes not soluble in water but soluble in alcohol can be used to sensitize collodio-silver halide emulsions (and from aqueous-alcoholic media, gelatino-silver halide emulsions also).

Thus examples might be cited among the azo- and bis-azo-dyes, the triphenylmethane dyes, the oxazine, thiazine and xanthene dyes, the alizarine dyes, the cyanines and phthalocyanines. Of these, the xanthene and cyanine dyes are of present practical importance.

Weak adsorption to silver halides, generally because of excess of solubilizing groups, accounts for relatively poor sensitizing in many cases, while the desensitizing tendency of many dyes, particularly for the normal blue-violet image, is another factor. However, from their preliminary survey, Eder and Valenta were able to deduce that (a) the sensitizing spectrum is, in general, closely related to the absorption spectrum of the solution—usually aqueous or alcoholic, and (b) the wave-length of maximum sensitizing is normally displaced toward the longer wave-lengths when compared with the maximum of absorption in such solutions. They considered that if it were possible to measure the spectral absorption of the dyed silver halide, i.e., of the adsorbed dye, it would be found to coincide with the spectral sensitivity conferred by the dye. This has been demonstrated recently by Natanson[3] for two xanthene dyes, *viz.*, erythrosine and phloxine, and for the cyanine dyes by Leermakers, Carroll and Staud.[4] We may regard the distribution of spectral sensitivity as determined by the absorption spectrum of the dye plus the effect of adsorption, i.e., as equivalent to the *adsorption spectrum*,[5] although there are dyes which are adsorbed to silver halide, but which do not sensitize, or only very weakly.

This wave-length shift observed by Eder and Valenta for many classes of dyes was noted by A. von Hübl[6] and others for the cyanine dyes upon their introduction. A group of *isocyanines*



(I)

having various substituents was examined by Sheppard,[7] measurements being made of both the absorption spectrum of solutions and of the sensitizing of silver bromide. The group *R* attached to the nitrogens was either ethyl or methyl, and the substituents in the quinoline nuclei were, respectively, halogen, alkyl or alkyloxy-groups. With all these isocyanines the sensitizing bands reproduced, but with displacement, the absorption bands of the solutions. The sensitizing maxima were shifted some 10 to 15 mμ relative to the absorptions in alcohol. These dyes showed a well-marked band of maximum absorption toward the red, and an adjacent band, often appearing only as a shoulder, of shorter wave-length. It was convenient to term these the α- and β-bands.

A similar study was made a decade later by (Miss) F. M. Hamer[8] of a group of isomeric isocyanines. The parent body was the 1 : 1′dimethyl-isocyanine, (cf. formula (I) and three different groups were substituted, positions as noted in Table I, which summarizes the observations.

The behavior with *acetyl amino-* variously substituted, and (by hydrolysis of the acetyl amino-derivatives) with the isologous amino-derivatives, was entirely similar. The two bands were displaced in sensitizing by much the same intervals from the bands in alcoholic solution. As previously found by Sheppard, the shift of the α-

TABLE I. *1 : 1′-Dimethyl-isocyanine iodides.*

DERIVATIVE	λ_{max} ETHYL ALCOHOL	λ_{max} SENSITIZING	Δ
Parent substance	β 520 mμ α 555	β 535 mμ α 585	15 mμ 30
5-cinnamoyl-amino	β 525 α 560	β 545 α 585	20 25
6-cinnamoyl-amino	β 535 α 575	β 550 α 600	15 25
7-cinnamoyl-amino	β — α —	β 540 α 595	— —
6′-cinnamoyl-amino	β 530 α 570	β 545 α 595	15 25
7′-cinnamoyl-amino	β 520 α 560	β 540 α 590	20 30

[3] S. Natanson, Nature **140**, 197 (1937).
[4] J. A. Leermakers, B. H. Carroll and C. J. Staud, J. Chem. Phys. **5**, 875 (1937).
[5] S. E. Sheppard, R. H. Lambert and R. D. Walker, J. Chem. Phys. **7**, 265 (1939).
[6] A. von Hübl, *Eder's Jahrbuch für Photochemie* (W. Knapp, Halle, 1905), p. 183.
[7] S. E. Sheppard, Phot. J. **32**, 300 (1908).
[8] F. M. Hamer, Phot. J. **62**, 8 (1922).

TABLE II.

Class of Dye	Example	Absorption Maxima in mμ: Alcohol	Absorption Maxima in mμ: Water	Sensitizing Maxima in mμ	Δ
I Pseudocyanine	1 : 1′-diethyl	β 490 α 522	485 522	β — α (570)	(48*)
II Isocyanine	1 : 1′-diethyl	β 515 α 555	520 555	β 525 α 570	10 15
III Cyanine	1 : 1′-diethyl	β — α 590	540 590	β — α 590 –610	~20
IV Carbocyanine	1 : 1′-diethyl-2 : 2′ (Pinacyanol)	γ 520 β 563 α 610	512 543 600	γ 540 β 587 α 637	20 24 27
V Di-cyanine (2 : 4′carbo)	1 : 1′-diethyl	β 613 α 660	580 650	β 638 α 690	26 30
VI 4 : 4′-carbocyanine	1 : 1′-diethyl-(Kryptocyanine)	β [640] α 715	625 710	β [690] α 750	35
VII Thiacarbo	2 : 2′-diethyl	β 525 α 562	510 558	β 535 α 585	10 23
VIII Thiacyanine	2 : 2′-diethyl	β 380 } 45 α 425 }	380 } 40 420 }	β overlaps with α natural sensitivity	
IX Thia-pseudo	1 : 2′-diethyl	β 450 } 40 α 490 }	— —	β — α [530]*	[40]*
X Thia-iso	1′ : 2-diethyl	β — α 508	470 } 30 500 }	β — α 525	25
XI Indo-carbo	1 : 1′-diethyl 3 : 3-3′ : 3′-tetramethyl	β 508 } 42 α 550 }	505 } 35 540 }	β 530 } 50 α 580 }	22 30
XII Oxa-carbo	2 : 2′-diethyl	β 450 } 35 α 485 }	450 } 30 480 }	β — α 510	25

* In the cases marked with an asterisk, in which the displacement is abnormally large (for this series), the dyes have the ψ-cyanine configuration. It is probable that in these cases the long wave sensitizing band is *not* the displaced molecular absorption band, but a new nematic phase discussed later. In fact, the wave-length 570 mμ assigned to the ψ-cyanine I is in close agreement with that observed later by Jelley and by Scheibe, references 39 and 40.

band was somewhat greater than that of the β-band. Miss Hamer further remarks "With each of these isocyanines the α-sensitizing maximum is situated 20 to 35 mμ nearer to the red end of the spectrum than is the corresponding absorption band [in alcohol]. It is clear from Table I that without exception, in the sensitizing curves of the three series, this crest in the cases of the 6-, 7-, and 6′-substituted isocyanines lies nearer to the red than in the equivalent 5- and 7′-substituted dyes. It is interesting that the correlation between the sensitizing and absorption spectra [in alcohol] is so close that an identical relationship exists in the latter case."

Somewhat later, O. Bloch and (Miss) F. M. Hamer[9] showed that the same correlation held for a whole range of differing classes of cyanine dyes, as exemplified in Table II.

[9] O. Bloch and F. M. Hamer, Phot. J. **68**, 22 (1928).

This correlation of sensitizing bands with absorption bands in solution is completely borne out by the many investigations on the preparation of sensitizing dyes of the cyanine groups. Particularly clear examples are given by L. G. S. Brooker and his collaborators in their "Studies in the cyanine dye series: IV"[10] in respect of the theretofore unknown thiazolo-2′-cyanines and oxazole cyanines synthesized by them. They found that the order in which these dyes fall is the same, whether absorption or sensitizing maxima are considered with respect to increasing wave-length; the difference between corresponding absorption [in alcohol] and sensitizing maxima varies from 42.5 mμ to 52.5 mμ. (It is notable that this is considerably larger than the displacements observed by Bloch and Hamer for other types of cyanine dyes (cf., Table II).)

Finally, in this connection, we would refer to the notable correspondence of sensitizing with absorption (in solution) shown in any series of cyanine dyes which are homologous merely by increase of the number of methine groups linking the two key nitrogen atoms. Table III shows this for the benzothiacyanines.[11]

Certain features of the quantitative relations involved, as between the wave-length of maximum absorption (or of sensitization) and the number of methine groups, and again between the extinction, or band-strength of the absorption maximum, and the number of methines, will be considered in a later section, and more fully elsewhere. It suffices at this stage to weight the argument further for correlation of absorption and sensitizing.

Significance of the Displacement

The foregoing evidence establishes a correlation or correspondence (a) directly between the absorption spectrum of the dye molecule in solution and its sensitizing spectrum and (b) inferentially between these and the molecular structure. We may conclude, therefore, that in this type of optical sensitizing the dye is molecularly engaged, indeed unimolecularly. Before ad-

[10] L. G. S. Brooker, G. H. Keyes and F. L. White, J. Am. Chem. Soc. **57**, 2492 (1935).

[11] J. Eggert, Veröff. wiss. Zentral Lab. phot. abt. Afga **3**, 136 (1933); **4**, 161 (1935). Also, W. Dieterle and W. Zeh, IX Congrès Int. Phot. Sci. Appl., (Paris, 1935), p. 188.

vancing to consider another type of sensitizing (and absorption) in which the dye is plurimolecularly engaged, we may ask, what is the significance of the shift observed between absorption and sensitizing? Here we may note that it was early shown by Sheppard[7] that the absorption of the isocyanine dyes was displaced to longer wave-lengths in different solvents as the refractive index of the solvent increased (Kundt's rule). Later investigations have shown this to be general for the cyanine dyes. In the particular case, the measurements show that referring the displacement to a solution in a given solvent, e.g., alcohol, is quite arbitrary, and that one could select a solvent of higher refractive index for which the displacement of the sensitizing maximum would be zero or negative.

Actually, any such displacement should be referred to the absorption spectrum of the dye in the gaseous state, and at low pressures. The displacement, referred to alcoholic solution, has no particular significance; referred to the absorption in the gas—it has significance in regard to the *adsorption energy* which will be discussed later. The "field effects" evident in the case of organic solvents have no direct bearing upon the problem of optical sensitizing. They are indirectly important, in regard to the relation between molecular constitution and absorption of light, and thereby, indirectly, for the problem of sensitizing.

A qualitative observation is pertinent in this connection. A cyanine dye, adsorbed to silver halide, gives a reflection spectrum corresponding to the sensitizing spectrum (and correspondingly displaced from that of the gaseous state). This surface was immersed in a liquid of high refractive index, which is also a solvent for the dye (e.g., methylene iodide $N_D = 1.74$). Although the solvent wets the layer of adsorbed dye, the reflection spectrum remains unchanged, and shows no displacement in the direction of the absorption shown by a solution of the dye in the same solvent. The unilateral adherence to the silver halide continues to dominate the response to light. As noted in a later discussion of solvent effects, they are of two kinds, one of nonpolar, the other of polar solvents. The effect of adsorption to silver halide seems more closely affiliated with the latter.

Structure of Absorption Bands of Cyanine Dyes

The band structure of the cyanine dyes in the visible region* has a definite affiliation (since it comprises a polymethine chain) with that of the polyenes which have been rather fully analyzed by Hausser and collaborators.[12] As with these, the wave-length of maximum absorption is progressively shifted to longer wave-lengths with increase in the number of methine groups, and—up to a limit—the extinction coefficient of this band also increases progressively. It is decisively modified, however, by the nitrogen-containing heterocyclic nuclei (quinolines, thiazoles, oxazoles, etc.) whose nitrogen atoms form the termini of the polymethine chains. Some of these characteristics are well illustrated by a transition form, the *merocyanines*, which have been synthesized by L. G. S. Brooker.[13]

The Absorption Band Structure of Cyanine Dyes

Numerous examples have been given of the absorption bands of various cyanine dyes in ethyl alcohol, and sometimes in aqueous solution. We desired to compare the bands observed in solution and on adsorption with those in the gaseous state, in order to evaluate if possible the "field" effects of the environment. It was found possible to sublime a number of cyanine dyes at low pressures and recondense them, either to thin mirrors or to thicker films. But when a tube length (optical path) sufficient to allow spectrographic or spectrophotometric measurements was used, decomposition (cracking or pyrolysis) oc-

Table III.

Dye	Number of Methines	Absorption Maximum in Mμ	Sensitizing Maximum in Mμ	Δ
2 : 2′benzothiacyanine	1	422	450	28
Benzothiacarbocyanine	3	557	595	38
Benzothia-di-carbocyanine	5	650	695	45
Benzothia-tri-carbocyanine	7	740	810	70
Benzothia-tetra-carbocyanine	9	845	915	70
Benzothia-penta-carbocyanine	11	960	1020	60

* Including the near ultraviolet and infra-red regions.

[12] K. W. Hausser, Zeits. f. tech. Physik **15**, 10 (1934).

[13] U. S. Patents 2,177,401–2–3.

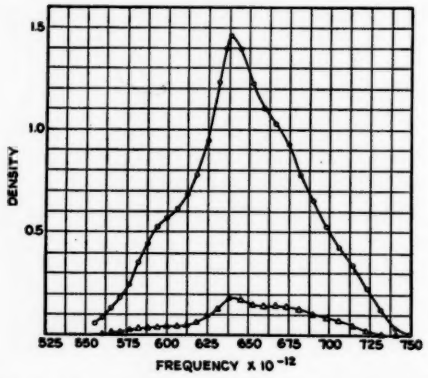
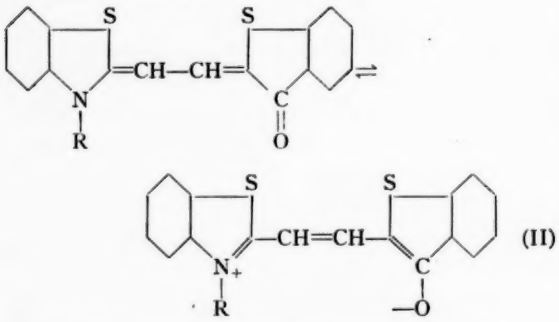


FIG. 1. Merocyanine vapor at △228°C, ⊙250°C.

curred sufficient to obscure the spectral absorption. In discussion Dr. L. G. S. Brooker suggested a trial of merocyanines. These dyes are not ionized, and it was found possible to volatilize them with less decomposition. The most satisfactory results so far have been obtained with 2(3-ethyl-2(3)(benzothiazolylidene-ethylidene)-2(3)thionaphthenone. This dye may be regarded as a mesomer or resonance hybrid between the extreme states (formulae II)



(II)

A fuller account of the investigation (particularly in regard to the effect of solvents) will be published* separately but the following results are pertinent to the present argument.

The maximum absorption in the gaseous state (at 228°C and 250°C, pressures (total) *ca.* 0.1μ) occurs at a frequency of 638 fresnels (10^{12} sec.$^{-1}$) or at wave-length 4700A, with other bands indicated at frequencies of 594, 674, and 714 fresnels. These are shown in Figs. 1 and 2.

Optical densities only are given and not extinctions, because the partial pressure of the dye was not known. In Fig. 2 the absorption band at 250°C has been analyzed into a series of bands by use of distribution functions of the form $y=k\exp(-x^2/2a^2)$ as introduced by V. Henri.[14] Here k is the ordinate at the center (maximum, for $x=0$) and $2a$ is the breadth at 0.606 k. A similar treatment of the band in hexane solution is shown in Fig. 3.

Examination of the bands in various nonpolar and polar solvents showed:

(a) The band system is displaced in solution, compared to the gaseous state, as a whole. There is little or no change of the intervals between the absorption bands.

(b) In nonpolar solvents the wave-length of the band of maximum absorption is a linear function of the refractive index N_D of the medium—including the gaseous state.

(c) In polar solvents there are increased displacements of this band toward longer wave-lengths, which are not a simple function of the dielectric constant nor of the dipole moment.

The close affinity of the band systems in different solvents and in the gaseous state is shown by Fig. 4, in which are plotted the ratios of the intensity (extinction or density) of the frequency of maximum absorption to the intensities (or extinctions) of the other bands. It will be seen also that the spectra in the gaseous state and in nonpolar solvents (hexane, carbon tetrachloride) have in common an extra band on the long wave side of the band of maximum intensity which is not present in polar solvents. The absorption appeared to follow Beer's law in all the solutions tested. The frequency differences between the

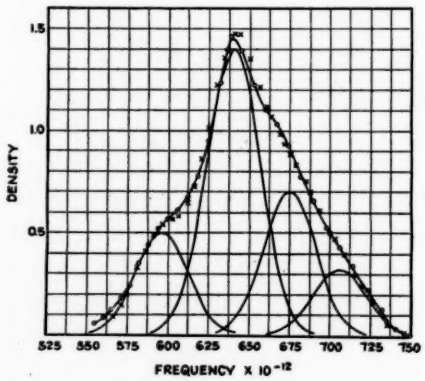


FIG. 2. Merocyanine vapor at 250°C. ⊙=observed, X=calculated.

* With P. T. Newsome.

[14] V. Henri, Physik. Zeits. **14**, 516 (1913).

sub-bands for different solvents are tabulated in Table IV.

The frequency or wave number difference varies somewhat according to the solvent. It also decreases, for the same solvent with the frequency of the band, but tends to a constant value as the frequency increases. Few of the cyanine dyes have been shown to present as many sub-bands in the visible region as this merocyanine so that less data as to the frequency interval (for any particular dye) are available. However, the values can be obtained for two or three readily discernible sub-bands for a large number of cyanine dyes of different classes, and some of these are presented in Table V.

In the case of the simple polyenes—notably the diphenyl polyenes—of increasing number of methine linkages[12] a Δ-value of about 1590 cm^{-1} (and, subordinately, one of 1240 cm^{-1}) is observed; this is very fairly constant, for the bands of the same polyene, and diminishes, but only slightly, as the number n of methines in the chain increases. These differences correspond to rotational-vibrational quanta in the infra-red spectrum, and, as further shown by Hausser and his collaborators[12] proper to the ethylene linkage >CH=CH<. Thus, the system of bands in the near ultraviolet and the visible regions results

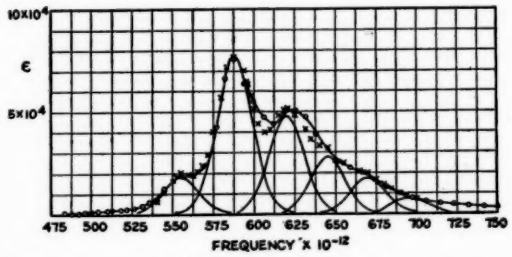


FIG. 3. Merocyanine in hexane (9.28×10^{-6} mole/liter). $y=ke^{-.000437x^2}$, ⊙ = observed, X = calculated.

from a fundamental electronic transition, on which are superposed vibrational quanta, and the spectra are represented by the term-formula

$$\nu=\nu_0+l\nu'+m\nu'',$$

where $\nu'=1590$ cm^{-1}

$\nu''=1240$ cm^{-1}

and l and m = (0, 1, 2, 3, etc.).

Hausser notes that with increasing n there is a small but steady fall of the vibrational frequency

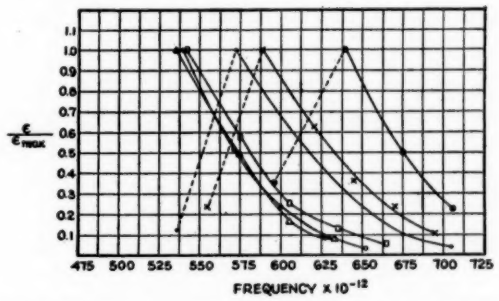


FIG. 4. Merocyanine (1.35×10^{-5} mole/liter). X = hexane, · = CCl_4, □ = quinoline, △ = methylene iodide, ⊙ = methyl alcohol, ⊗ = vapor-static.

in the infra-red of the >CH=CH< linkage, as shown by Raman spectra, and this, accompanying as it does a displacement of maximum absorption toward longer wave-lengths shows that the corresponding electron loosening is accompanied by a loosening of the >CH=CH< bond.

The absorption spectra of the *lin*-benzenes show in the ultraviolet a series of bands separated by constant frequency differences. The principal wave number differences run:

DIPHENYL POLYENE	ANTHRACENE	NAPHTHALENE	BENZENE
1600 cm^{-1}	1430 cm^{-1}	1360 cm^{-1}	900 cm^{-1}

and we may perhaps regard this as concomitant with the progressive relaxation and degeneracy of the >CH=CH< linkage in passing to benzene. In the case of the cyanine dyes, in which we have heterocyclic nuclei (also with fused-on aromatic, hydroaromatic, etc., nuclei), frequency differences of all these degrees of magnitude appear, corresponding to a greater complication of the superposition of vibrational frequencies on a fundamental electronic transition. Also it must be remembered that the alternating chain comprises as well as >CH=CH< linkages (one or two) terminal >C=N− or >C=O bonds.

In recent theories of the spectra of conjugated double bond systems, the long wave band of maximum extinction is ascribed to a fundamental electron transition proper to "the most extended form of the molecule" (R. S. Mulliken[15]) or to the greatest range in the molecule of an electron transfer (and reciprocal positive charge transfer) (L. Pauling[16]). In one analysis, "the higher frequency bands correspond to more crumpled or

[15] R. S. Mulliken, J. Chem. Phys. **7**, 570 (1939).
[16] L. Pauling, Proc. Nat. Acad. Sci. **25**, 577 (1939).

bent forms of the molecules,"[15] in another to abbreviated or incomplete stages of the potential electron transfer—involving intermediate structures, "since the electrons have vanishingly small probability of jumping from one end to the other of a large molecule, but make the transit instead by moving from atom to adjacent atom."[16] This appears to be consonant with the superposition of a vibrational frequency possible to a given link. This conception of the nature of the auxiliary higher frequency bands (in the visible) of the cyanine dyes is strengthened by the absorption spectra of the isomeric carbocyanines (Table VI and Fig. 5).

TABLE IV.

Band	Δ cm⁻¹ Vapor	Hexane	Carbon tetrachloride	Methyl Alcohol	Quinoline	Methylene Iodide
I						
II	1450	1100	1170	—	—	—
III	1200	1067	1250	1100	1120	1330
IV	1040	833	1180	950	1000	1000
V		832	1000	917	1000	950
VI		834	1000	833	1000	900

The structure of the absorption bands is very similar, as pointed out by Mills and Odams.[17] The primary bands at lowest frequency are displaced to lower frequency with increase of the conjugated chain between the N atoms extended in this case by change of position of the nuclei. Each band has a well marked secondary maximum at about 1250 cm^{-1} from the primary, and a feeble tertiary (at least in IIa and III) at a further interval of 1300 to 1400 cm^{-1}. The spectrum of the 2 : 4′-carbocyanine is intermediate between that of *pinacyanol* and the *kryptocyanine*, in accord with its structure.* The strengthening of the primary bands, of lowest frequency, on increase of chain length, is correlated with the fading of the tertiary band. Very noteworthy is the fact pointed out by Mills and Odams that the secondary maxima of kryptocyanine and the 2 : 4′-carbocyanine lie nearly at the same frequencies as the primary maxima of the 2 : 4′-carbocyanine and pinacyanol, and that the tertiary maximum of the 2 : 4′-carbocyanine corresponds closely in position with the secondary maximum of pinacyanol. Similar correspondences are observable with the simple polyenes, less evident with homologous cyanines (increasing $-(CH)_n$), by reason of the very rapid intensification of strength of the chief band (of lowest frequency) and associated weakening of the auxiliary bands.

The question as to their nature is of considerable importance in the case of aqueous solutions, when these auxiliary higher frequency bands are strengthened. The comparison of the spectra of a homologous series of cyanines, differing only by the number n of methine groups between the nitrogens also supports this relation of the subsidiary bands to the band of greatest absorption. In Fig. 6 are shown the absorption curves of such a series (alcoholic solution). The presence of subordinate bands is indicated by the asymmetry on the short wave-length side, which appears to increase considerably for the dye with greatest λ_m.

Ascription of the higher frequency subsidiary bands to stereoisomers (obtained by rotation about links of the polymethine chain) as suggested by G. Scheibe[18] does not appear satisfactory, in part for reasons noted shortly. Also in such a case, the absolute transit of the electron need not be affected, but possibly the transition moment, and thereby the relative intensity (extinction coefficient). This might be regarded as

TABLE V.

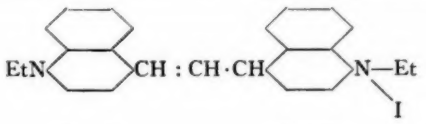
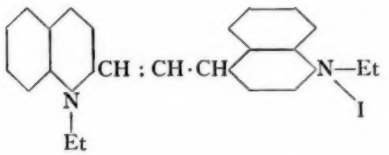
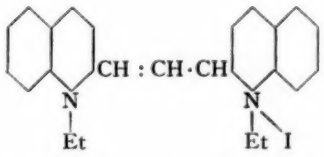
Δ cm⁻¹ Isocyanines				
Sheppard	Bloch and Hamer	Carbocyanines	Indocarbocyanines	Oxacarbocyanines
1420	1200	1440	1520	1600
1200	1300	1400		
1170	1340	1200		
1100	1240	1650		
940		1200		
860				

[17] W. H. Mills and R. C. Odams, J. Chem. Soc. II, **125**, 1913 (1924).

* As shown by L. G. S. Brooker, reference 27, this conclusion does not hold in general without important qualifications.

[18] G. Scheibe, Zeits. f. angew. Chem. **52**, 631 (1939).

TABLE VI.

I (Kryptocyanine)	1 : 1′-diethyl- 4 : 4′ carbocyanine iodide	
IIa (Diocyanine)	1 : 1′-diethyl- 2 : 4′ carbocyanine iodide	
III (Pinacyanol)	1 : 1′-diethyl- 2 : 2′ carbocyanine iodide	

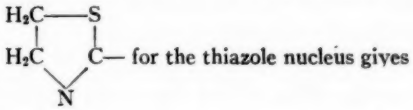
shown by comparison of the spectra of *trans*- and *cis*-stilbene.[19] The spectrum of the *trans*- or extended form shows much higher extinctions at about the same frequency maxima as for the *cis*-form. However, the *cis*-form cannot take a coplanar form, and this factor must be taken into account in the reduction of the intensity of absorption.

A steric factor of very great importance both for absorption and for sensitizing is planarity of the dye molecule. This will be discussed more fully in its place, but cases where the terminal nuclei would approach too closely (nearer than 3.1∼3.5A) in the extended coplanar form for this to be viable involve a modification (or enhancement) of the nuclear effects which should be mentioned here, as perhaps not a primary but in any case a secondary nuclear influence.

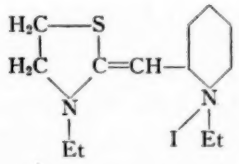
This apart, it is clear that the band system of the cyanine dyes is not simply compounded of the electron transitions possible to a polyene chain extended between two nitrogen atoms, with possible rotational-vibrational frequencies superposed. Thus the spectrum is decisively modified by the character of the nuclei containing these nitrogen atoms (e.g., quinoline, thiazole, etc.) by substitution in these nuclei and again by the alkyl or other extranuclear groups attached to the nitrogens.

Beside the general mass of evidence of correlations of this sort, very direct evidence is given by the work of Fisher and Hamer[20] on the correlation of frequencies of maximum absorption with changes in chemical constitution with series of symmetrical cyanines, and again by that of Beilenson, Fisher and Hamer[21] on unsymmetrical cyanines. In the latter work it was shown that the position of the wave-length of maximum absorption for unsymmetrical cyanines (different terminal nuclei) is given approximately by the mean wave-length between the wave-lengths λ_m and λ_m' for the corresponding symmetrical cyanine.*

Another striking example of the nuclear auxochromic effect is given by L. G. S. Brooker's work on dyes derived from 2-methyl thiazoline.[22] The substitution of the partly saturated nucleus, i.e.,



for the thiazole nucleus gives



(III)

1′3-diethyl-thiazolino-2′-pyridocyanine iodide. In

[19] Cf., A. Smakala and A. Wasserman, Zeits. f. physik. Chemie **A155**, 353 (1931).

[20] N. I. Fisher and F. M. Hamer, Proc. Roy. Soc. **A154**, 703 (1936).

[21] B. Beilenson, N. I. Fisher and F. M. Hamer, Proc. Roy. Soc. **A163**, 138 (1937).

* But with notable exceptions, depending on the relative basicity of the nuclei.

[22] L. G. S. Brooker, J. Am. Chem. Soc. **58**, 662 (1936).

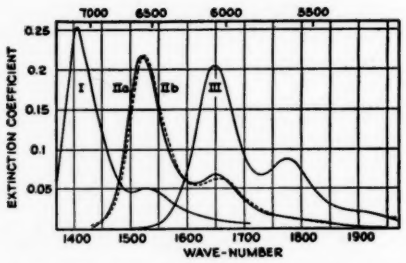


FIG. 5. Spectral absorption of isomeric carbocyanines.

addition, from *2-iodoquinoline ethiodide* was prepared *1',3-diethyl-thiazolino-2'-cyanine iodide* and from *2-iodo-β-naphthoquinoline ethiodide*, the *1',3-diethyl-5',6'-benzo-thiazolino-2'-cyanine iodide.* Comparing absorption spectra, with those of the corresponding dyes derived from thiazole, Brooker remarks that "the effect of partly saturating the thiazole nucleus is to shift the absorption bands nearer the blue"—in the language of the chromophore theory, a hyposochromic effect. Conversely, the fusion of benzo-groups on to the pyridyl nucleus displaced the absorption toward the red. The effects of saturation outside the primary resonance chain are very striking.

The basic chemical and structural cause for these effects has only become gradually apparent with the progress of the theory of the relation of color to constitution. The chromophore *cum* auxochrome classification led over by way of the *quinonoid* aperçu,** to the recognition of chains of conjugated double bonds as the most potent chromophores. Even so the influence of groups not comprised in the effective chain was not clear. An important advance was registered, so far as a classificatory basis was concerned, by W. König's systematization of the polymethine dyes[23] in terms of a fundamental *mesochrome*—the conjugated chain, and attached *perichrome* portions of the molecule (which correspond, but more generally, to the terminal nuclei just discussed). The phenomenon of *halochromy*—radical shift of absorption centers on salt formation by ammonium, or generally "*onium*" cation formation, or (*aci*halochromy) by anion formation, of the *peri*chrome groups—was then a further pointer, but tended to short-circuit the explanation in terms of the concomitant but inessential ionization.

The important advance made in the application of the resonance theory to color by Bury[24] helped show the function of the auxochromes in a chromophore—by increasing the resonance states possible to the chromophore, but did not at first clarify the functions of the perichrome. The essential conceptions were developed in respect to the general field by Chr. Ingold[25] in a review of *Electronic Theories of Organic Reactions* and were touched upon, in specific relation to sensitizing and desensitizing dyes, by J. D. Kendall.[26] In our opinion, the latter's contradistinction between sensitizers and desensitizers is inadequate, if not incorrect, but it will suffice to quote his definitions to show that, qualitatively, the function of the terminal or perichrome groups was definitely apprehended. Proceeding from Ingold's review, he states, "On voit alors que tous les sensibilisateurs contiènnent ce que l'on est convenu d'appeler 'système neutralisé' dans lequel les centres anionöides (*donneurs* d'électrons) et cationöides (accepteurs d'électrons) de la molécule sont liés par une chaine non saturée et dans lesquels le polarisation induite conduit, par déplacement électromérique, à la formation de dipoles comme indiqué par la formule générale:

$$\overset{+}{\alpha}-(C=C)_n-C=\overset{-}{\beta},$$

ou $n=0, 1, 2, 3$, etc., α l'un des éléments O, S, Se, N, etc., comme dans les composants anionöides, et β l'un des mémes éléments, comme dans les composés cationöides."

It is unnecessary to discuss here Kendall's speculations concerning the "dicationoid" character of the end-groups in the case of desensitizing dyes. The important thing is the recognition that the relative electron donating or accepting capacities of the perichrome (end-groups) can change the color (deepen or lighten) with the same intermediate chain. This is readily perceived on comparing the symmetrical cyanine

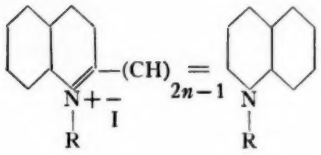
** And in some respects, in spite of it.

[23] W. König, J. prakt. Chem. (2) **112**, 1 (1926).

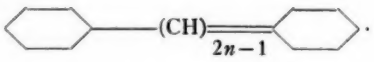
[24] C. R. Bury, J. Am. Chem. Soc. **57**, 2115 (1935).

[25] C. K. Ingold, Chem. Rev. **15**, 225 (1934).

[26] J. D. Kendall, IX. Congrès Internat. Phot. Sci. et Appl. (Paris, 1935), p. 227.



with the corresponding symmetrical diphenyl polyene



The subject is very fully and clearly covered in forthcoming papers by L. G. S. Brooker and co-workers.[27] Brooker shows how varying basicity of heterocyclic end-groups can account for deviation from Fisher, Hamer and Beilenson's[21] relation (q. v.) between symmetrical and asymmetrical cyanines. He shows, on the one hand, how inequality in this respect of the terminal nuclei, making for inequality of energy for mesomeric states, reduces the resonance energy, while on the other hand, in symmetrical combinations, approach to saturation in the terminal groups also reduces the resonance. What Brooker terms "the resonance mobility" seems to be related to the degeneracy of double bond linkages in conjugated chains, as discussed previously (p. 101).

A direct picture of the reciprocal interaction of the polymethine chain and the attached nuclei is given by comparing respective bands both in the visible and in the ultraviolet for dyes and for the component nucleus. This should be done for (a) the colored dye at high *p*H, (b) the dye decolorized at low *p*H, (c) the nucleus, as free heterocyclic base, at high *p*H, and (d) the same as quaternary ammonium salt, at low *p*H. Specific examples of these are given in the curves of Fig. 7 for the isocyanine orthochrome *T*. A fuller and more comprehensive study, covering cyanines with various nuclei and various chain-lengths will be presented separately, but this much is clear. Whether in the colored form (a) or the colorless form (b), the ultraviolet bands, which might be more specifically ascribed to the nuclei (c) and (d), show great modification by their attachment to the polymethine chain—or to each other through this chain. This is the correlate to the fact that the absorption bands in the visible are determined both in position and intensity not only by the polymethine chain between the nitrogens, but by the terminal nuclei. This factor is further modified, in certain cases, by steric influences now discussed.

Stereoisomerism and Steric Impedance

At present the most important steric factor of the dye molecule for optical sensitizing, and to some extent for color itself, appears to be the planarity of the molecule. In the first paper[5] we presented evidence that the cyanine—and probably also the xanthene—sensitizing dyes are adsorbed to silver halide crystals "edge-on," as closely packed as the size and shape of the molecule, and the underlying lattice allowed. A planar form of the molecule was assumed, and a large degree of parallel piling of the edge-on flat molecules, and values calculated on these assumptions were found to agree best with the experimental data. This is the more important, because present theories of the color of dyes postulate a planar form for fullest development of the resonance conditions, and therewith a vibration or transition polarized in the plane of the molecule. A molecule extended in three dimensions might have distinct oscillations relatively resolved along three axes, x, y, and z. For definitely planar molecules, this reduces to two (x, y). Again, structures such as the polyenes, and also the cyanines, in extended planar forms behave, insofar as concerns the bands controlled by the polymethine chain, practically as *linear* molecules or oscillators, for light absorbed in these bands. On the other hand, for ultraviolet light, absorbed in the upper partial bands dom-

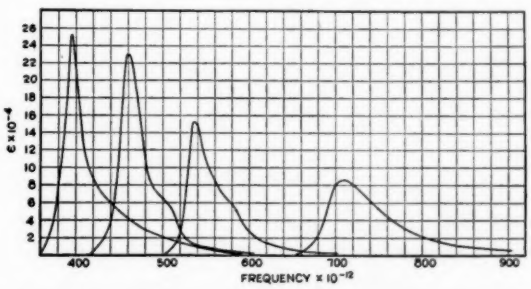


FIG. 6. Spectral absorption of homologous cyanine series. (Data from L. G. S. Brooker, R. H. Sprague, C. P. Smyth and G. L. Lewis, J. Am. Chem. Soc. 62, 1116 (1940).)

[27] To be published (Part I., L. G. S. Brooker, R. H. Sprague, C. P. Smyth and G. L. Lewis, J. Am. Chem. Soc. 62, 1116 (1940)).

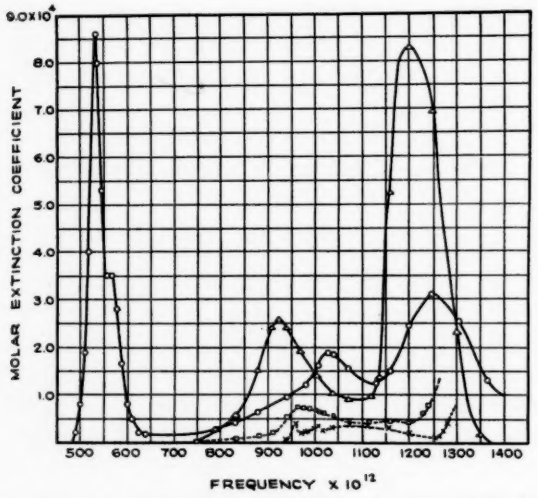
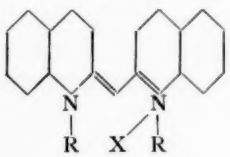


FIG. 7. Spectral absorption of orthochrome T bromide (free and acidified) and of quinoline (free and acidified). Solvent—methyl alcohol, acid—sulfamic acid. X = quinoline 3.9×10^{-4} M, ⊙ = orthochrome T bromide 2.3×10^{-5} M, ⊡ = quinoline (acidified) 2.0×10^{-4} M, △ = orthochrome T bromide (acidified) 2.3×10^{-5} M.

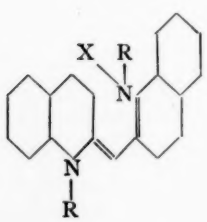
inated by the nuclei, both x and y bands are feasible—but are not of direct importance for color and color sensitizing. But planar isomers (where obtainable by permissible rotations of 180° about a >CH−CH< linkage) could have importance for color—if only through reduction of extinctions by change in the transition moment, as pointed by G. Scheibe[18] in this connection, and more generally indicated by Förster.[28] Again, if we consider two planar stereoisomers such as

Formula IV

cis



trans

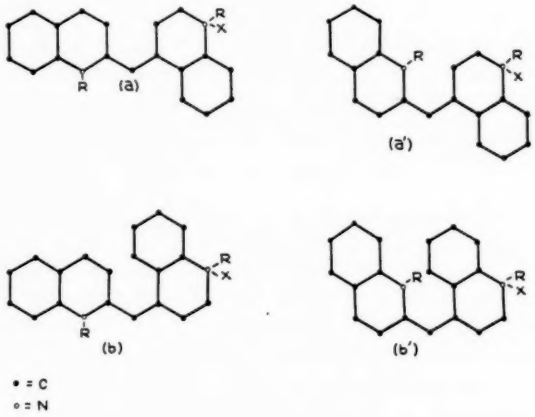


it seems probable that the opposite position of the N atoms of the resonance chain—which in a sense are alternately seats of the positive charge—might interfere with and reduce edge-on adsorption to silver halide, as well as reduce the extinction, for the aforementioned reason.

The specific delimitation of the intervention of possible stereoisomeric forms of polymethine dyes in optical sensitizing requires extensive and precise comparison of: (i) the steric forms of series of dye molecules, (ii) the absorption spectra in solution, (iii) their adsorption to silver halide, and (iv) their photographic sensitizing and photolytic abilities.

Some progress has been made on this program, but attention must now be directed to a very important and possibly crucial steric factor affecting sensitizing.

The isocyanine (formula I) could be ascribed four stereoisomeric planar configurations:



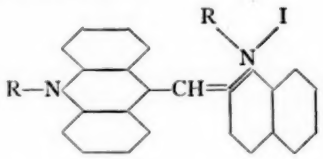
Of these, a and a' (the latter obtained from a by rotation about a chain linkage) give an unbonded approach of =C− to −C= or of =C− to −N= of about 2.7A, assuming bond angles of 120° and interatomic distance =C···C− of 1.4A. By assuming sufficient strain to make the angle $\beta = 140°$, and the chain links to be extended from 1.4A to 1.54A, this distance can be expanded to 3.4A sufficient to allow planarity of the molecule. On the other hand, in configurations b and b', the approach between the two proximate nuclei becomes 1.6A, and, as in the case of diphenyl benzene[29] a fully planar form is not viable, it being necessary to assume rotations

[28] T. Förster, Zeits. f. Elektrochemie **45**, 548, 752 (1939).

[29] J. M. Robertson, J. Chem. Soc. 232 (1939).

of the nuclei out of the plane, of the order of 40° to 50° to secure a clearance of 3.1 to 3.2A. The diminution of resonance involved makes this configuration unlikely compared with the *a*, *a'* form. The dye shows two well-marked narrow bands in solution and is a sensitizer. It is probable that the form *a*, with the N atoms on the same side of the molecule (or cation) is dominant on adsorption to silver halides.

Miss F. M. Hamer[30] on attempting to prepare cyanine dyes from 5-methyl acridine obtained (but only when using water as a solvent) by condensation between 5-methyl acridine and 2-iodoquinoline methiodide or ethiodide dyes of the formula (V)



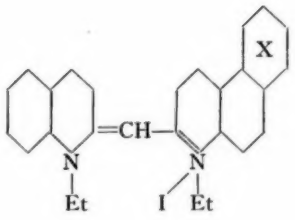
analogous to the isocyanine just described.

Miss Hamer remarks that "although ··· derived from the simple isocyanine, (v. sup.) their absorption bands are of a different character, showing one broad band, with an ill-defined maximum ··· in approximately the same position as that of the chief band of the parent isocyanine, instead of showing the expected shift toward the red ··· the new dyes are *entirely devoid* of sensitizing properties."

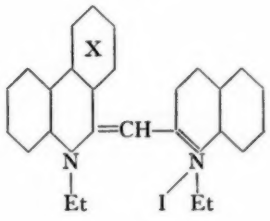
Now on laying out the corresponding planar stereoisomeric configurations to those just presented for the simple isocyanine it is readily seen that in no case is a coplanar configuration possible in which the minimum approach of unbonded nuclear carbons or carbon-nitrogen atoms is less than 1.6A—assuming normal bond angles and atomic distances.

This correlation of weakened and diffused absorption and of lack of sensitizing with steric impedance in the molecule of coplanarity is further established by the following examples:

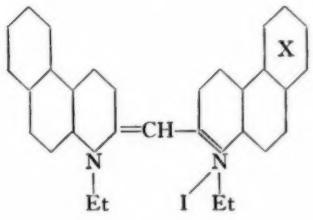
L. G. S. Brooker and G. H. Keyes[31] have prepared and described certain isomeric cyanine dyes derived from α-naphthoquinoline and from phenanthridine, respectively, and having the formulae:



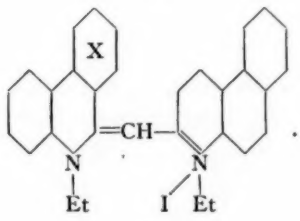
(A)



(B)



(C)



(D)

They observed that dyes A=XVIc and C=XVIIIc were strong sensitizers, while B=XVIIc and D=XIXc did not sensitize. These substances were investigated later by J. A. Leermakers, B. H. Carroll and C. J. Staud[4] who confirmed the sensitizing results. But in addition they noted adsorption to gelatino-silver bromide grains of *all* the isomers and "pronounced absorption maxima (by reflection)."

In recent work on these compounds, we have observed, first, that the *rate* of adsorption to silver bromide from aqueous gelatin solution of dyes B=XVII and D=XIX is very much lower than for dyes A=XVI and C=XVIII. This is brought out by the curves of Fig. 8, for amounts of dye adsorbed after 18 hours' tumbling of an

[30] F. M. Hamer, J. Chem. Soc. 995 (1930).
[31] L. G. S. Brooker and G. H. Keyes, J. Am. Chem. Soc. 58, 659 (1936).

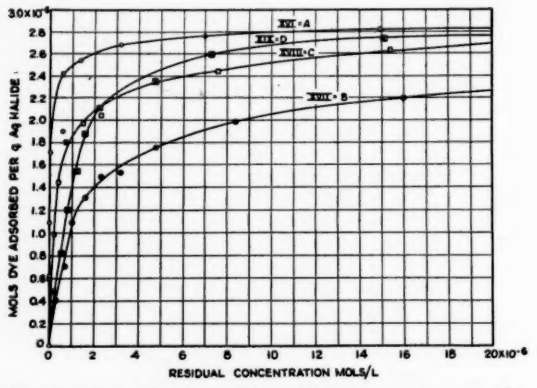


FIG. 8. Rate of adsorption of benzocyanines to silver halide. ○ = dye XVI = *A*, □ = dye XVIII = *C*, ⊗ = dye XVII = *B*, ⊠ = dye XIX = *D*.

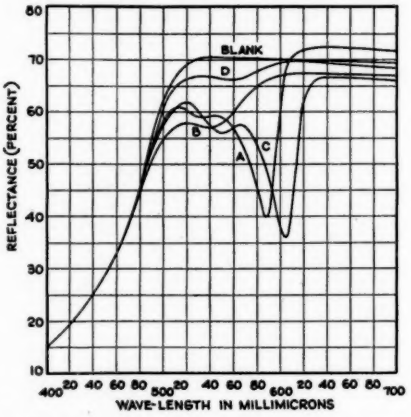


FIG. 9. Reflection spectra of benzocyanines adsorbed to silver bromide.

emulsion at 40°C. It will be seen that at low concentrations of dye, the adsorption densities of the two nonsensitizing dyes B = XVII and D = XIX, are very much lower. The ultimate adsorption densities reached—in terms of moles dye adsorbed per gram of silver halide, appeared somewhat lower (about 10 percent) for the nonsensitizing isomers, in presence of gelatin.

While the low rate of adsorption—and especially for normal sensitizing concentrations—might account in part for ineffective sensitizing, it does not appear a sufficient explanation.

In the reflection spectra of the dyes adsorbed to silver bromide, while all four dyes show bands of not very great intensity corresponding to their absorption bands in alcoholic solutions (cf. Fig. 9), the two sensitizing isomers, A = XVI and C = XVIII, show two very sharp strong bands for which there are no equivalents in alcoholic solution. These bands are caused by "aggregate" absorption, of a mesomorphic form of the dyes discussed in a following communication.

A sharper difference in behavior between the isomers was brought out by experiments on the adsorption of the dyes to precipitated silver bromide and on the photolysis of the dyed silver halide. This was determined, in the presence of a halogen acceptor, by methods previously described.[32] The adsorptions in this case—absence of gelatin—advanced to close to the theoretical ratios at saturation for the areas subtended by the molecules edge-on[5] as shown in Table VII.

[32] S. E. Sheppard, R. H. Lambert and R. D. Walker, J. Chem. Phys. 7, 426 (1939).

But the photolytic yields were entirely different. From the curves of Fig. 10 it can be concluded that these nonsensitizing dyes, although adequately adsorbed, produce *no photolysis* for light absorbed by the dye. This is obviously in complete agreement with their failure to sensitize photographically.

This complex of differences we believe is accounted for by steric impedance of a planar form for two of the isomers.

On constructing the formulae for these isomers, using correct bond angles and interatomic distances, it can be seen that complete planarity of the molecule is not possible in case B without bringing the 3,4-benzo-radical within a distance of *ca.* 1.7A of the opposite quinoline group, and in case *D* of the opposite β-naphthoquinoline. This condition is not improved by 180° rotation of a nuclear group about a =CH−CH linkage to an alternate planar stereoisomer. The same condition obtains for *1,1'-diethyl-3,4,3',4'-dibenzo-2,2'-carbocyanine iodide* (having *three* methines between *two* phenanthridine nuclei) which was found also not to sensitize, but does not for

TABLE VII. *Surface of* AgBr *Ppt.* = 25,000±10 cm^{-2}/g.

Dye	Area per Dye Molecule	Moles Dye Ads./g AgBr	
		Calculated	Observed
A	56A²	7.94×10^{-6}	7.92×10^{-6}
B	52	7.37	7.34
C	60	6.88	6.60
D	56	7.37	7.14

1,1'-diethyl-5,6,5',6'-dibenzo-2,2'-carbocyanine iodide, which sensitizes strongly, (having three methines between two β-naphthoquinoline nuclei) and also not for *1,1'-diethyl-3,4,5',6'-dibenzo-2,2'-carbocyanine iodide* having three methine groups between one β-naphthoquinoline and one phenanthridine nucleus) and which definitely sensitizes.

Again, on inspecting the absorption spectra (in alcoholic solution), it can be seen that the band strengths for the planar isomers are much greater than for the nonplanar isomers, the molecular extinctions at the wave-lengths of maximum absorption being about twice.

It was suggested by Leermakers, Carroll and Staud[4] that "the group *X* in dyes *B* and *D* interferes with the proper exchange of energy between dye and silver halide, because of some steric effect in the molecule." This steric effect we consider to be again the inhibition of planarity.

It is probable that attention would have been devoted earlier to this, at any rate more examples would have been uncovered in the cyanine dyes, had not investigation moved largely on to the synthesis of cyanine dyes of increased chain length. With this the likelihood of steric interference between the nuclei becomes reduced—although, up to a point, is still able to affect the *number* of viable planar isomers of the same dye. Considering this chain length then, it is of interest to glance at the so-called *apocyanines*, in which the nuclei are directly bonded, without intervention of the methine group. The formula

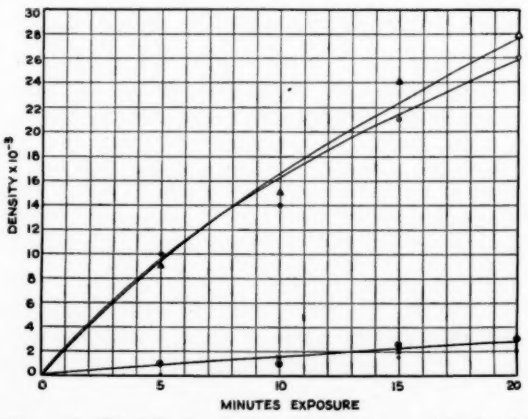
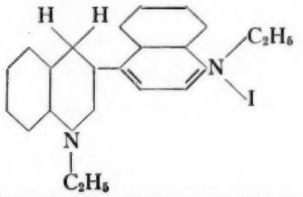


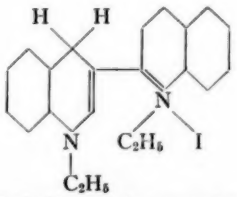
FIG. 10. Photolysis of silver bromide sensitized with benzocyanine dyes. ·=blank, *X*=*B* and ⊗=*D* nonsensitizers, ⊙=*A* and △=*C* sensitizers. Wave-length of illumination=546 mμ.

assigned to the so-called *erythro-apocyanine*, and *xantho-apocyanine*, respectively, by Mills and Ordish[33] are shown in planar form:



(VI)

1 : 1'-Diethyl(erythro)apocyanine iodide



(VII)

1 : 1'-Diethyl(xantho)apocyanine iodide.

In both cases the planar form involves an unbonded approach of about 2.8A, which, however, can hardly be increased by any strain, because the same interval occurs on both sides of the orthogonal axis between the bonded nuclei. As with diphenyl and its derivatives[29] a coaxial but to some extent nonplanar form seems probable.

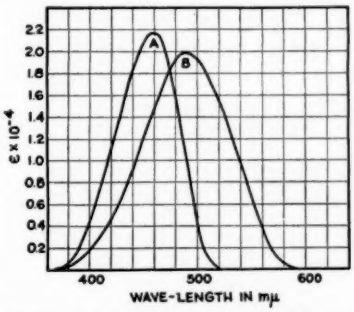


FIG. 11. Spectral absorption of apocyanines (in alcohol) *A*—xantho-apocyanine, *B*—erythro-apocyanine (recalculated from data of O. Bloch and F. M. Hamer, (Phot. J. **59**, 21 (1928)).

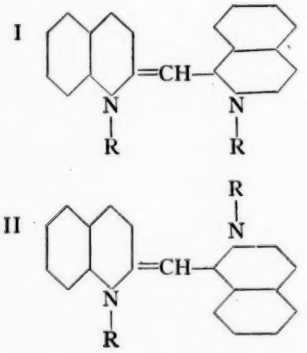
This seems partly borne out by the observations that the band strengths of the apocyanines are quite low (Fig. 11), while the optical sensitizing is stated by Hamer (and Bloch)[9] to be slight by bathing, and to be practically nil when the dyes are put in the emulsion.

As further examples of cyanine dyes in which

[33] W. H. Mills and H. G. Ordish, J. Chem. Soc. Part I, 81 (1928).

a greater or less degree of steric impedance to planarity might be expected, we may mention cyanine dyes containing an isoquinoline nucleus, described as 2 : 1′ and 4 : 1′-cyanines* by (Miss) N. I. Fisher and (Miss) F. M. Hamer, who first prepared them.[34] It is interesting to note these authors' general comment, "although the new groups of cyanines containing an *iso*-quinoline nucleus do comprise dyes possessing color-sensitizing properties, their action is, in general, definitely weaker than that of the corresponding compounds which contain a quinoline in place of the *iso*-quinoline nucleus."

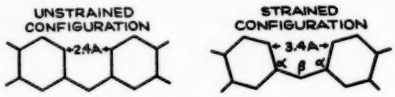
2 : 1′-cyanines



(VIII)

In I and II are shown two stereoisomeric planar isomers of the 2 : 1′-cyanine conformation. These are viable provided a degree of strain on the bond angle of the methine be tolerated of the same order as postulated for ψ-cyanines. Two further planar stereoisomers, obtained from the foregoing by rotation through 180° of the isoquinoline nucleus about the $-CH-1'$ bond are excluded, because bringing the nuclei too close. The colors of the 2 : 1′-cyanines are said to be optically indistinguishable from those of the corresponding ψ-cyanines (q.v.). We have not yet had the opportunity to ascertain whether their solutions show a *negative temperature coefficient* of absorption, as we find with ψ-cyanines. From the strain necessary to give planar molecules we should expect this. The case is similar for the 4 : 1′-cyanines also described by Fisher and Hamer in the same paper. However, the relatively broad (diffuse) absorption bands reported for these 2 : 1′- and 4 : 1′-cyanines cannot be solely ascribed to steric influences, and there may well be direct nuclear or perichrome effects.

In further confirmation of the influence of planarity on the light absorption, we may note that dyes having the ψ-cyanine structure can only possess a planar configuration by allowing a considerable strain on certain bond angles of the mesochrome (see examples).



When the absorption of either aqueous or alcoholic solutions of such a dye (1′ : 1′-dimethylψ-cyanine) was measured at different temperatures, it was observed that the extinction diminished on raising the temperature, and increased again to its original value on cooling. This phenomenon will be dealt with more fully in a later publication. But such a negative temperature coefficient of absorption points to relaxation of the strain on the bond angles by increase of the rotational-vibrational energy of the molecule.

There seems to be something fundamental and essential for optical sensitizing in the edge-on adsorption of *planar* dye molecules—or ions—oriented in a definite azimuth to the (cubic) silver halide lattice. In a letter to *Nature* (June 22, 1940) we suggested that the essential orientation was perpendicular "to a (110) surface plane of the crystal" and parallel to internal (110) planes. Mr. A. P. H. Trivelli, of these Laboratories, has pointed out to us that the indices given are incorrect for an octahedral plane: They should be (111), as given in our first paper on adsorption.[5] If now congruence of the plane of the adsorbed molecule, or ion with a (111) plane is of prime importance for optical sensitizing, the planes of the adsorbed dye ions should be tilted at about 70° to a surface (111) plane. Mr. Trivelli has suggested that the *intersections* of (110) planes with (111) planes may be the critical loci—in any case, planar *congruence* of the adsorbed dye ions with intersecting (110) planes would require a tilt of 35°; this seems at present somewhat improbable, from adsorption density considerations, but further experiments are planned.

We will outline here a mechanism of optical sensitizing which will be dealt with more fully in

* Nomenclature according to Hamer and Fisher.
[34] F. M. Hamer, J. Chem. Soc. 1905 (1934).

a later communication. The conditions just noted imply a high degree of *congruence* of the absorbed dye molecule with the lattice structure of the ionic crystal. We suppose that the electrons of the filled (strongly shielded) zone[35] (atomistically of the bromide ions of a 111 plane) have a vector of energy E_a in the 111 plane. This component may undergo coupling with a parallel component of the resonance energy of the adsorbed dye polarized in the same azimuth. This would correspond to the adsorption energy of the dye, and to the displacement of the light absorption of the adsorbed dye molecule to lower frequencies in the free gaseous state (Table VIII inf.). This in effect supposes this coupling to proceed prior to absorption of light and hence in the ground levels. On the other hand, in accordance with conceptions advanced by de Boer,[36] this displacement may be regarded as coming from a different adsorption energy of the excited state on irradiation. Quantitatively, this does not appear to affect the energy balance, and one may for the present consider it (the adsorption energy) as an enhancement of the previously assumed kinetic energy component for an electron of the filled level, which thereby becomes $E_a+h\Delta\nu$. On absorption of a quantum $h\nu$, of light by the adsorbed dye, the coupling advances to the taking over of this whole increment of electronic transition, polarized in the same azimuth, by such a partially coupled electron of the filled zone; the total kinetic energy of this electron in the given plane becomes $E_a+h\Delta\nu+h\nu \gtrless h\nu_0$, where $h\nu_0$ is the energy necessary to raise an electron to the conductance zone of the crystal. The condition becomes similar in certain respects to that deduced by A. von Hippel[37] for the dielectric breakdown in alkali halide crystals, in which it was shown that in these (cubic) crystals the electrons prefer to move through the lattice in specific planes.

Essentially, what is postulated is a sufficiently large coplanar component of the otherwise strongly shielded electronic energy of the filled zone to make up, by coupling with a photon absorbed by the dye molecule, the energy necessary to free the electron. We suggest that *optical sensitizing* is limited to precisely the most favorable conditions for this. The relation of the suggested mechanism to so-called metastable levels and disturbed lattice points will be discussed in a later paper, as also the interesting suggestion of J. Eggert[38] that in ordinary latent image formation by ultra- and blue-violet light the more perfect lattice regions are acting as "optical sensitizers" for perturbed or imperfect regions.

The foregoing conclusions, derived for planar cyanine dye molecules and "molecular" sensitizing, can be shown to be perfectly consistent with "aggregate" sensitizing.[4] It has been shown that the adsorbed molecular aggregates are parallel-piled ensembles of planar dye ions adsorbed edge-on.[5] And it has been shown by E. E. Jelley[39] and by G. Scheibe[40] that filaments or packets of these parallel-piled ions constitute an initially nematic phase. The new- and strongly shielded-electronic transition (shown in absorption, in fluorescence, and in optical sensitizing) proper to these aggregates is polarized in a direction perpendicular to the plane of the component molecular ions, i.e., parallel to the axis of the packet or micelle—and therefore to a 111 plane of the crystal. The coupling of this with the previously indicated azimuthal component of the electronic energy of the filled zone could proceed therefore as for an adsorbed molecule. In this connection it is significant that we find that nonplanar cyanine dyes do not aggregate to the "nematic" state, do not give the new band.

A more complete theory of the "aggregate" state, and of the associated spectral absorption, and the essential part played by intermolecular water molecules will be dealt with in a following paper.

The considerations advanced are equally applicable to acid sensitizing dyes, e.g., erythrosine, of the xanthene type.[5] These also are adsorbed edgewise to an octahedral surface though of silver ions. The molecules of these dyes are also planar, indeed, *a fortiori* because having a fused polycyclic structure. So far as azimuthal coupling is concerned, it is immaterial whether adsorption is

[35] Cf., H. H. Wilson, Proc. Roy. Soc. **A133**, 458 (1931); **A134**, 277 (1931).
[36] A. de Boer, *Electron Emission and Adsorption Phenomena* (Cambridge, 1935), p. 198.
[37] A. von Hippel, Zeits. f. Physik **67**, 707; **68**, 309 (1931); **75**, 145 (1932).
[38] J. Eggert, Trans. Faraday Soc. **34**, 892 (1938).
[39] E. E. Jelley, Nature **139**, 631 (1937).
[40] G. Scheibe, Kolloid Zeits. **82**, 1 (1938).

TABLE VIII.

No.	Dye	Maximum Absorption of Vapor Fresnels Calc.	Maximum Sensitizing Fresnels	$\Delta\nu$ (Vapor-sens.) Fresnels	$h\Delta\nu$ ev
VIIIb	Orthochrome T bromide	562	517	45	0.183
VIIId	Orthochrome T acetate	—	517	—	—
IVb	3,3′-diethyl-9-methyl thiacarbocyanine bromide	561	509	52	0.215
IVa	3,3′-diethyl-9-methyl thiacarbocyanine chloride	563	509	54	0.223
XXIIc	3,3′-diethyl-9-iso-amyl thiacarbocyanine iodide	556	500	56	0.227
Xe	3,3′-diethyl-thiacarbocyanine p-toluene sulfonate	563	509	54	0.223
Ia	3,3′-diethyl-9-methyl-4,5,4,5′-dibenzo-thiacarbocyanine chloride	519? Vapor absorption measured	484	35	0.143
XXIII	2(3-ethyl-2(3)benzothiazolylidene-ethylidene)-2(3)thionaphthenone	706	556	150	0.618
		674	509	165	0.681
		638	469	169	0.696

to a bromide ion or to a silver ion layer. The electrons of the filled zone are, energetically speaking, under constraints from both bromine and silver.

The "adsorption energy" contribution to sensitizing will be discussed more fully in a following communication, but certain values may be given to indicate the order of magnitude it can attain.

In Table VIII it can be seen that the contribution is considerable, and may even attain a magnitude sufficient in itself, to supply energy for photolysis (and sensitizing). However, as will be noted in the later publication, this does not appear to be necessarily the case.

Extra-Perichrome Substituents

The effects of substituents—alkyl, alkyl-oxy, aryl-, halogens, etc.—in the perichrome (or nuclear) portion of the cyanine molecule can be largely accounted for by electromeric and steric influences. More uncertain is the origin of the effects, e.g., of alkyl groups attached either to the terminal N atoms of the mesochrome, or substituted for hydrogen in the (central) methine group of the chain. This uncertainty is due in the first place to the relatively meager and incomplete data. Presumptively systematic investigations on these constitutional factors were initiated by W. H. Mills and W. J. Pope[41] in 1920 for the isocyanines and the carbocyanines. Unfortunately, no quantitative data were given for the *absorption spectra* in solution, while the sensitizing spectra were obtained by bathing only, and in an admittedly rather incomplete and tentative manner.

As regards alkyls attached to the N atoms, for isocyanines the conclusions in Table IX are recorded.

The dipropyl and dibutyl spectra showed progressively more marked minima between sensitizing bands, and it might be concluded that sensitizing strength diminished with molecular weight of the alkyl radicals. However, with the carbocyanines[42] a somewhat different picture was obtained (Table X).

In these series there are two examples of weak sensitizing by the 1 : 1′-dimethyl dye, and, in the second increase to the diethyl and dipropyl, and then a decline. It should be noted here that in both (a) and (a_1) considerable general (blue-violet) *desensitizing* was apparent in the spectra recorded.

This weak sensitizing by the methiodides of the carbocyanines was confirmed for the *thiacarbocyanines* by F. M. Hamer[43] and also for the 1 : 1′-di-alkyl-4 : 4′-diphenyl carbocyanine by Hamer, Heilbron, Reade and Walls.[44] The dye *4 : 4′-diphenyl-1 : 1′-dimethyl quinazo-carbocyanine iodide* was stated to be "practically devoid of either sensitizing or desensitizing properties," but as no comparison appears to have been made with the *ethiodide*, it cannot be concluded with certainty if this was due to the methyl groups. Similar results, however, were obtained with styryl derivatives of quaternary salts, except that here the *methiodide* (with a broad absorption band at 530 mμ) "exerts considerable desensitizing power." This desensitizing power was still

TABLE IX.

Dye	Sensitizing Maxima	Degree of Sensitizing
1 : 1′-dimethyl-6-methyl (iodide)	585 mμ : 535 mμ	Strong xxx
1 : 1′-diethyl-6-methyl (iodide)	586 mμ : 535 mμ	Strong xxx
1 : 1′-dipropyl-6-methyl (iodide)	591 mμ : 535 mμ	Somewhat less xx
1 : 1′-dibutyl-6-methyl (iodide)	581 mμ : 535 mμ	Somewhat less xx

[41] W. H. Mills and W. J. Pope, Phot. J. **60**, 183 (1920).

[42] W. H. Mills and W. J. Pope, Phot. J. **60**, 253 (1920).

[43] F. M. Hamer, J. Chem. Soc. 2061 (1929).

[44] F. M. Hamer, I. M. Heilbron, J. H. Reade and H. N. Walls, J. Chem. Soc. Part I, 251 (1932).

further increased in *4-phenyl-2-p-hydroxystyryl quinoline methiodide*. It is possible that the effect of methyl groups is not so much a simple lessening of optical sensitizing, as an enhancement of *desensitizing*.[5] Whether this, or a quasi-steric effect from a hydrogen of the methyl on the basicity of the N atom, similar to that observed by J. F. J. Dippy[45] on the dissociation constants of mono-carboxylic acids, is responsible for the peculiar behavior of the methiodides cannot be affirmed at present, since comparative adsorption and photolytic data are lacking.

The effect of simple substituents (alkyls, aryls, halogens) in the central methine of the chain is in some ways more important and comprehensible. Ground was broken in this direction by (Miss) F. M. Hamer in an important paper in 1928.[46] The effect of alkyls is in general slightly hypsochromic as may be seen from Table XI from Brooker and White.[47]

TABLE X.

DYE	SENSITIZING MAXIMA	DEGREE OF SENSITIZING
a. 1 : 1′-dimethyl-iodide	*ca*. 580 mμ	Very weak
b. 1 : 1′-diethyl-iodide	*ca*. 540∼580 mμ	Very strong
a_1. 1 : 1′-dimethyl 6 : 6′-dimethyl-iodide	550, 650–690 mμ	Rather weak
b_1. 1 : 1′-diethyl 6 : 6′-dimethyl-iodide	550, 650–690 mμ	Strong
c_1. 1 : 1′-di *n*-propyl 6 : 6′-dimethyl-iodide	650–700	Strong
d_1. 1 : 1′-di *n*-butyl 6 : 6′-dimethyl iodide	570∼640	Less than c_1

TABLE XI.

CHAIN SUBSTITUENT	ABSORPTION MAXIMUM IN Mμ		
	BENZOTHIO-CARBOCYANINE	4 : 5 : 4′ : 5′-DIBENZO-BENZO-THIOCARBO-CYANINE	6 : 7 : 6′ : 7′-DIBENZO-BENZO-THIOCARBO-CYANINE
Unsubstituted	558	590	593
Methyl	543	566	570
Ethyl	548	568	575
n-Propyl	548		
n-Butyl	548		
n-Amyl	548		
Iso-amyl	548		
Benzyl	550		
Phenoxy-methyl	553		
Phenyl	560	593	595

Together with this hypsochromic effect in absorption there is stated to be increased effectiveness in sensitizing, although this appears to be largely a matter of improved green sensitizing concomitant with the changed absorption, and therefore of no (other) theoretical interest. The central substitution of another quaternary base (as in neocyanine[22]) giving cross-conjugation and *two* systems of absorption and sensitizing does not require discussion in the present connection.

[45] J. F. J. Dippy, Chem. Rev. **25**, 151 (1939).

[46] F. M. Hamer, J. Chem. Soc. Part II, 3160 (1928).

[47] L. G. S. Brooker and F. L. White, J. Am. Chem. Soc. **57**, 2480 (1935).

TABLE VIII.

No.	Dye	Maximum Absorption of Vapor Fresnels Calc.	Maximum Sensitizing Fresnels	$\Delta\nu$ (Vapor-sens.) Fresnels	$h\Delta\nu$ ev
VIIIb	Orthochrome T bromide	562	517	45	0.183
VIIId	Orthochrome T acetate	—	517	—	—
IVb	3,3′-diethyl-9-methyl thiacarbocyanine bromide	561	509	52	0.215
IVa	3,3′-diethyl-9-methyl thiacarbocyanine chloride	563	509	54	0.223
XXIIc	3,3′-diethyl-9-iso-amyl thiacarbocyanine iodide	556	500	56	0.227
Xe	3,3′-diethyl-thiacarbocyanine p-toluene sulfonate	563	509	54	0.223
Ia	3,3′-diethyl-9-methyl-4,5,4,5′-dibenzo-thiacarbocyanine chloride	519? Vapor absorption measured	484	35	0.143
XXIII	2(3-ethyl-2(3)benzothiazolylidene-ethylidene)-2(3)thionaphthenone	706	556	150	0.618
		674	509	165	0.681
		638	469	169	0.696

to a bromide ion or to a silver ion layer. The electrons of the filled zone are, energetically speaking, under constraints from both bromine and silver.

The "adsorption energy" contribution to sensitizing will be discussed more fully in a following communication, but certain values may be given to indicate the order of magnitude it can attain.

In Table VIII it can be seen that the contribution is considerable, and may even attain a magnitude sufficient in itself, to supply energy for photolysis (and sensitizing). However, as will be noted in the later publication, this does not appear to be necessarily the case.

Extra-Perichrome Substituents

The effects of substituents—alkyl, alkyl-oxy, aryl-, halogens, etc.—in the perichrome (or nuclear) portion of the cyanine molecule can be largely accounted for by electromeric and steric influences. More uncertain is the origin of the effects, e.g., of alkyl groups attached either to the terminal N atoms of the mesochrome, or substituted for hydrogen in the (central) methine group of the chain. This uncertainty is due in the first place to the relatively meager and incomplete data. Presumptively systematic investigations on these constitutional factors were initiated by W. H. Mills and W. J. Pope[41] in 1920 for the isocyanines and the carbocyanines. Unfortunately, no quantitative data were given for the *absorption spectra* in solution, while the sensitizing spectra were obtained by bathing only, and in an admittedly rather incomplete and tentative manner.

As regards alkyls attached to the N atoms, for isocyanines the conclusions in Table IX are recorded.

The dipropyl and dibutyl spectra showed progressively more marked minima between sensitizing bands, and it might be concluded that sensitizing strength diminished with molecular weight of the alkyl radicals. However, with the carbocyanines[42] a somewhat different picture was obtained (Table X).

In these series there are two examples of weak sensitizing by the 1 : 1′-dimethyl dye, and, in the second increase to the diethyl and dipropyl, and then a decline. It should be noted here that in both (a) and (a_1) considerable general (blue-violet) *desensitizing* was apparent in the spectra recorded.

This weak sensitizing by the methiodides of the carbocyanines was confirmed for the *thiacarbocyanines* by F. M. Hamer[43] and also for the 1 : 1′-di-alkyl-4 : 4′-diphenyl carbocyanine by Hamer, Heilbron, Reade and Walls.[44] The dye *4 : 4′-diphenyl-1 : 1′-dimethyl quinazo-carbocyanine iodide* was stated to be "practically devoid of either sensitizing or desensitizing properties," but as no comparison appears to have been made with the *ethiodide*, it cannot be concluded with certainty if this was due to the methyl groups. Similar results, however, were obtained with styryl derivatives of quaternary salts, except that here the *methiodide* (with a broad absorption band at 530 mμ) "exerts considerable desensitizing power." This desensitizing power was still

TABLE IX.

Dye	Sensitizing Maxima	Degree of Sensitizing
1 : 1′-dimethyl-6-methyl (iodide)	585 mμ : 535 mμ	Strong xxx
1 : 1′-diethyl-6-methyl (iodide)	586 mμ : 535 mμ	Strong xxx
1 : 1′-dipropyl-6-methyl (iodide)	591 mμ : 535 mμ	Somewhat less xx
1 : 1′-dibutyl-6-methyl (iodide)	581 mμ : 535 mμ	Somewhat less xx

[41] W. H. Mills and W. J. Pope, Phot. J. **60**, 183 (1920).

[42] W. H. Mills and W. J. Pope, Phot. J. **60**, 253 (1920).

[43] F. M. Hamer, J. Chem. Soc. 2061 (1929).

[44] F. M. Hamer, I. M. Heilbron, J. H. Reade and H. N. Walls, J. Chem. Soc. Part I, 251 (1932).

further increased in *4-phenyl-2-p-hydroxystyryl quinoline methiodide*. It is possible that the effect of methyl groups is not so much a simple lessening of optical sensitizing, as an enhancement of *desensitizing*.[5] Whether this, or a quasi-steric effect from a hydrogen of the methyl on the basicity of the N atom, similar to that observed by J. F. J. Dippy[45] on the dissociation constants of mono-carboxylic acids, is responsible for the peculiar behavior of the methiodides cannot be affirmed at present, since comparative adsorption and photolytic data are lacking.

The effect of simple substituents (alkyls, aryls, halogens) in the central methine of the chain is in some ways more important and comprehensible. Ground was broken in this direction by (Miss) F. M. Hamer in an important paper in 1928.[46] The effect of alkyls is in general slightly hypsochromic as may be seen from Table XI from Brooker and White.[47]

TABLE X.

DYE	SENSITIZING MAXIMA	DEGREE OF SENSITIZING
a. 1 : 1′-dimethyl-iodide	*ca*. 580 mμ	Very weak
b. 1 : 1′-diethyl-iodide	*ca*. 540∼580 mμ	Very strong
a_1. 1 : 1′-dimethyl 6 : 6′-dimethyl-iodide	550, 650–690 mμ	Rather weak
b_1. 1 : 1′-diethyl 6 : 6′-dimethyl-iodide	550, 650–690 mμ	Strong
c_1. 1 : 1′-di *n*-propyl 6 : 6′-dimethyl-iodide	650–700	Strong
d_1. 1 : 1′-di *n*-butyl 6 : 6′-dimethyl iodide	570∼640	Less than c_1

TABLE XI.

CHAIN SUBSTITUENT	ABSORPTION MAXIMUM IN Mμ		
	BENZOTHIO-CARBOCYANINE	4 : 5 : 4′ : 5′-DIBENZO-BENZO-THIOCARBO-CYANINE	6 : 7 : 6′ : 7′-DIBENZO-BENZO-THIOCARBO-CYANINE
Unsubstituted	558	590	593
Methyl	543	566	570
Ethyl	548	568	575
n-Propyl	548		
n-Butyl	548		
n-Amyl	548		
Iso-amyl	548		
Benzyl	550		
Phenoxy-methyl	553		
Phenyl	560	593	595

Together with this hypsochromic effect in absorption there is stated to be increased effectiveness in sensitizing, although this appears to be largely a matter of improved green sensitizing concomitant with the changed absorption, and therefore of no (other) theoretical interest. The central substitution of another quaternary base (as in neocyanine[22]) giving cross-conjugation and *two* systems of absorption and sensitizing does not require discussion in the present connection.

[45] J. F. J. Dippy, Chem. Rev. **25**, 151 (1939).

[46] F. M. Hamer, J. Chem. Soc. Part II, 3160 (1928).

[47] L. G. S. Brooker and F. L. White, J. Am. Chem. Soc. **57**, 2480 (1935).

The Thermal Reaction Between Hydrogen and Oxygen

II. The Third Explosion Limit

O. OLDENBERG AND H. S. SOMMERS, JR.
Research Laboratory of Physics, Harvard University, Cambridge, Massachusetts
(Received October 19, 1940)

The thermal high pressure explosion of the hydrogen-oxygen mixture ("third explosion limit"), occurring near atmospheric pressure at 560°C, was studied. In the p–T diagram, a section of the limiting curve was determined. The effect of a KCl coat on a Pyrex surface was observed. The investigation of the third explosion limit does not lend itself to a decision between the competing theories of the thermal hydrogen-oxygen reaction.

I. PROBLEM

HYDROGEN and oxygen when mixed at nearly atmospheric pressure in the stoichiometric proportion, at a temperature between 500° and 570°C, form water with a measurable rate. However, as discovered by Hinshelwood and his co-workers,[1] explosion occurs if, at a given temperature, the pressure is either lowered or raised (Fig. 1; below C or above D, respectively. In the present paper we are not concerned with the low pressure explosion limit A included for the sake of completeness in Fig. 1. Limit C is called the "second" explosion limit and limit D the "third.")

These limits and the behavior of the steady reactions approaching them are of great interest for the theory, which has been worked out in detail by Lewis and von Elbe.[2] In a preceding paper we dealt with some aspects of the second explosion limit which offer considerable theoretical difficulty. The present paper is concerned with the high pressure side of the steady reaction, that is, the third explosion limit. Because of the destructive effect this type of explosion has not been investigated in detail until recently.[3]

For the purpose of a spectroscopic investigation of free radicals, we started with the problem of producing the thermal reaction with a maximum steady rate, that is, working as close as possible to the third explosion limit. Some incidental explosions showed that, with our small vessels, the explosion at the third limit is gentle enough to be accessible to a careful investigation. It is not as destructive as the explosion of the mixture $2H_2+O_2$ at room temperature and atmospheric pressure started by a spark; this difference is doubtless caused by the difference in densities. During several hundred explosions in glass cylinders of about 360 cc at pressures between 47 and 77 cm and temperatures between 540 and 575°C, not a single vessel was broken. In a longer vessel (length 120 cm) with plane quartz windows, however, the explosion did shatter the windows.

II. APPARATUS AND PROCEDURE

The reaction was observed in both Pyrex and quartz vessels, most of them being cylindrical of 5 cm diameter and 12 cm length. The reaction vessel was shielded by wire screens. The vessel was evacuated with a mercury pump backed by a "Hyvac" oil pump or, instead, by a "Megavac" pump that gave a final pressure too low for electric discharges; by rinsing the vessel with hydrogen any residual impurities could be washed out. Pressures were observed with a mercury manometer. To prevent condensation of water vapor, the manometer as well as the inlet tubes of the reaction vessel could be heated.

The following gives the operating procedure: H_2 was admitted, the manometer heated, and finally, O_2 admitted and the reaction—steady

[1] See C. N. Hinshelwood and A. T. Williamson, *The Reaction between Oxygen and Hydrogen* (Oxford, Clarendon Press, 1934). Since the "third explosion limit," discussed in the present paper, is above the so-called "upper" explosion limit, we replace the term "upper" by "second" explosion limit.

[2] B. Lewis and G. von Elbe, J. Am. Chem. Soc. **59**, 656 (1937), and *Combustion, Flames, and Explosions of Gases* (Cambridge, University Press, 1938).

[3] O. Oldenberg and H. S. Sommers, Jr., J. Chem. Phys. **7**, 279 (1939).

rate or explosion—observed. (Practically no mercury vapor could diffuse from the manometer into the reaction vessel except the 1/1000 mm present at room temperature. It was checked that higher concentrations of mercury vapor did not measurably change the rate.)

The electric oven was carefully built for uniform temperature. This was adjusted, by extra coils in the top and bottom of the cylindrical oven, to ±2°C over the reaction vessel. The temperature was measured with a Chromel-Alumel thermocouple and potentiometer, the unit being calibrated against the melting point of a zinc sample furnished by the National Bureau of Standards.

For our first experiments on the rates of the thermal reaction, we used electrolytic gases dried over long columns of P_2O_5. Later, for the sake of convenience, we used tank gases. At the first trial with these the reaction differed strikingly in that it abruptly stopped far before completion, although the impurities as stated by the manufacturer (0.5 percent N_2, A, and CO_2) are known not to affect the reaction. Presumably the lubricating oil from the hydrogen compressor changed the reaction. Therefore we bubbled the tank hydrogen through a long layer of Nujol oil contained in a helical washing bottle, passed it through tightly packed glass wool cooled by solid CO_2, and dried it over P_2O_5. By this procedure we obtained the same reaction as with

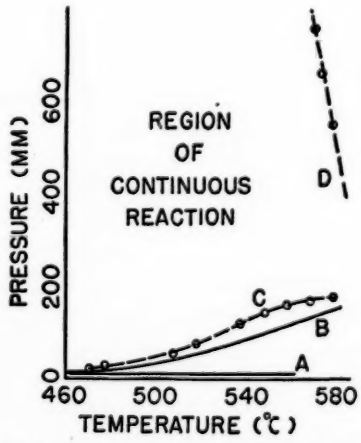


FIG. 1. Explosion limits of the mixture $2H_2+O_2$. *A*, lower limit, glass or quartz; *B*, second limit, glass or quartz; *C* second limit, KCl; *D*, third limit, KCl.

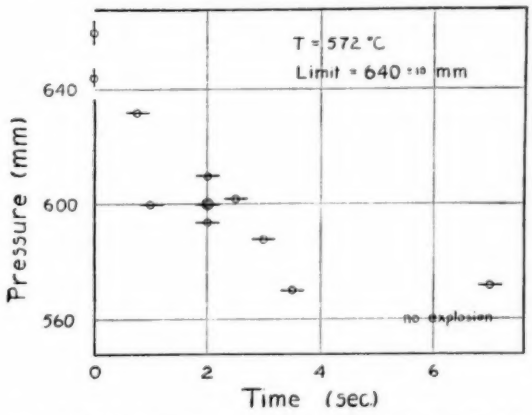


FIG. 2. High pressure explosion (third explosion limit); inhibition period as a function of pressure.

electrolytic gases.[4] The hydrogen was admitted before the oxygen, because in these regions, near the third explosion limit, admission in the reverse order produces an instantaneous explosion or very rapid combustion as soon as any H_2 is added to the O_2. This initial combustion is presumably a case of the explosion below the second (or "upper") limit. It is avoided when beginning with H_2, because for a given stoichiometric mixture twice the pressure of this gas is admitted, compared with O_2.

The explosion is observed as a flash in the reaction vessel and a clearly audible click. At the same moment the mercury manometer jumps and then suddenly drops to a lower level with a speed obviously limited by its inertia. The explosion is preceded by an inhibition period which, in an extreme case, may be as long as 95 seconds. This makes the determination of the limiting values of pressure and temperature rather indefinite, for at a given temperature, explosion eventually ensues for a certain range of pressures after varying inhibition periods. Since the higher the pressure, the shorter is the inhibition period, we defined the explosion limit as that p-T combination for which the inhibition period vanished. This definition led to the following procedure: At a given temperature the inhibition period was plotted as a function of the initial pressure and extrapolated back to zero inhibition. In the example of Fig. 2

[4] We appreciate the advice of Dr. G. B. Kistiakowsky. Garstang and Hinshelwood obtained the same rates with very pure gases and tank gases (Proc. Roy. Soc. London **A130**, 644 (1931)).

($T=572$°C) 640 mm is taken as the critical pressure. This type of curve was determined for various temperatures between 569 and 577°C. We could not well extend the temperature range any further, because for lower temperature the pressure exceeds one atmosphere—the highest pressure to be reached with our apparatus—while, on the other hand, for higher temperature initial burning commences with the first admission of O_2.

III. Pyrex Vessel Covered With KCl

According to Prettre's[5] results a KCl coating on the reaction vessel makes the reaction far more reproducible than it is in the uncoated vessel. Because of this, we investigated in particular the reaction in Pyrex vessels in which KCl was deposited from aqueous solution to form a thick visible layer. As discovered by Pease,[6] the KCl coating makes the reaction much slower. With the help of this coating and by taking our runs in rapid succession with only two minutes of pumping intervening, we were able to get satisfactory reproducibility over a long series of experiments.

With the procedure described above, the upper section of the explosion curve (D in Fig. 1) was obtained. As it is based on more than 150 runs, its accuracy is not spoiled by incidental errors. In particular the steep slope of the curve is sure to be real.

IV. Comparison with Theory

This branch of the explosion curve is satisfactorily explained by the theory for the thermal hydrogen-oxygen reaction of Lewis and von Elbe. In this theory, the details of which will not be discussed here, hydrogen atoms representing members of a chain reaction are removed by triple collisions:

$$H+O_2+M\rightarrow HO_2+M. \qquad (VI)$$

The radical HO_2 so formed is supposed to be rather inactive, even at temperatures around 550°C, so its fate depends largely upon the pressure. *At low pressure* it is supposed to diffuse to the wall where it is destroyed. This makes reaction (VI) chain breaking. At still lower pressures, however, triple collisions required for reaction (VI) get so rare that the chain breaking fails, and the increased yield of the individual chain explains the low pressure explosion (second or upper explosion limit). On the other hand, *at high pressures* diffusion to the wall is so slow that HO_2, in spite of its inactivity, has a chance to continue the chain by

$$HO_2+H_2\rightarrow H_2O+OH. \qquad (X)$$

The radical OH so produced easily continues the chain. The mathematical treatment of this reaction scheme explains the high pressure branch of the explosion curve described in the preceding section.

Can this satisfactory agreement be considered to be a strong argument in favor of the theory? We hardly believe that it can because it seems that other theories, too, may be capable of explaining the experimental curve. Prettre[5] successfully describes his measurements of reaction rates by the Bonhoeffer-Haber scheme which we shall not discuss in detail. At the first glance, this scheme may seem inferior to the one given by Lewis and von Elbe because the third explosion limit cannot be mathematically deduced. But this is no serious deficiency, since the mathematical treatment assumes fixed values of the rate constants as they apply at constant temperature. It is sufficient that the Bonhoeffer-Haber scheme gives a rate strongly increasing with the pressure. Internal heating, caused by the reaction itself, then explains the explosion. (The same heating process is certain to occur if the Lewis-von Elbe scheme is correct. Although their theory leads to an explosion even for an isothermal reaction, the authors state that internal heating would lead to an explosion for a lower pressure than given by their isothermal formula.)

In general, this seems to us a feature of any reaction scheme that describes a high order thermal reaction: *for high pressure* internal heating must cause the explosion. Furthermore, it follows that *with increasing temperature*, when all steps are speeded up, not so much of a pressure is needed to cause the explosion. This would explain the slope of the curve described (Fig. 1). We conclude that the investigation of the third

[5] M. Prettre, J. Chim. Phys. **33**, 189 (1936).
[6] R. N. Pease, J. Am. Chem. Soc. **52**, 5106 (1930).

explosion limit fails to give a stringent decision between the competing theories.

V. Change of Relative Concentration and Vessel Diameter

The theory of Lewis and von Elbe gives a definite equation showing how the third explosion limit should change with the partial pressures and the vessel diameter. Omitting the details, the final formula may be written

$$[H_2]d^2 = \text{constant}.$$

This "constant," expressed in terms of rate constants of the individual steps, depends only upon the temperature (Eq. (68) of Lewis and von Elbe). It follows that at constant temperature the limiting pressure should be independent of the O_2 concentration. In order to check this theoretical prediction we kept the H_2 pressure constant (40 cm) and varied the O_2 pressure. At 22.4 cm O_2 pressure the explosion occurred almost without inhibition, while at O_2 pressures of 4.1 or 12.2 cm no explosion occurred at all. Though this result seems to disagree with the theory, reconciliation of the theory with the experiments might be effected by assuming different efficiencies for the various types of molecules in the two effects: chain breaking in triple collisions and chain prolongating by slowing the diffusion of HO_2 to the walls. Therefore, the discrepancy is minor and does not furnish an argument against the theory.

The above formula predicts a considerable effect of vessel diameter. For comparison with the results given in Section III (vessel diameter 5 cm) we studied the explosion in a Pyrex vessel of 6.2 cm diameter, again covered with KCl. The result was that increasing the diameter definitely lowered the explosion temperature (at a pressure of 65 cm from 572 to 565°C). That is, it increased the tendency towards explosion. This result compares favorably with the theory. One obviously must not expect a quantitative agreement, because the theory is based on transportation of the chain carrier HO_2 to the wall by diffusion, while in the experiments uncontrollable *convection currents* are certain to add their effect.[7] The Bonhoeffer-Haber scheme, too, leads to such an effect of increasing vessel size favoring the explosion, since chain breaking at the wall becomes rarer in the larger vessel.

VI. Effect of Surface Conditions

The third explosion limit is also affected by the KCl coating. In the larger vessel (diameter 6.2 cm) at a pressure of 60 cm this limit is about 10°C lower in clean Pyrex than in KCl.

The following contrast between the *third* explosion limit (high pressure) and the *second* explosion limit (low pressure) is striking: at the *high pressure explosion* both effects of KCl go together, which is plausible (reduced rate and reduced chance for the *third* explosion) while, on the other hand, at the *low pressure explosion*, although KCl again reduces the rate, it *increases* the chance for the explosion (second or upper explosion, curves *B* and *C* of Fig. 1).

These apparently inconsistent effects of KCl at the low pressure side lead to theoretical difficulties discussed in a preceding paper. (These difficulties, in combination with other arguments, led us[8] to attribute the *second* explosion limit to the wall reaction and so assume it to be independent of the *volume reaction*. Both kinds of reaction are known to occur simultaneously, but their mutual relation, if there is any, is not well understood.) There is no such difficulty for the *third* explosion limit discussed in the present paper.

---

[7] The effect of convection currents has been discussed by D. A. Frank-Kamenetzky, Acta Physicochimica U. R. S. S. **10**, 365, (1939) and O. K. Rice, J. Chem. Phys. **8**, 727 (1940).

[8] O. Oldenberg and H. S. Sommers, Jr., J. Chem. Phys. **8**, 468 (1940).

Letters to the Editor

THIS section will accept reports of new work, provided these are terse and contain few figures, and especially few halftone cuts. The Editorial Board will not hold itself responsible for opinions expressed by the correspondents. Contributions to this section should not exceed 600 words in length and must reach the office of the Managing Editor not later than the 15th of the month preceding that of the issue in which the letter is to appear. No proof will be sent to the authors. The usual publication charge ($3.00 per page) will not be made and no reprints will be furnished free.

The Vibrational Energy Levels and Specific Heat of Ethylene

E. J. BURCIK, EUGENE H. EYSTER AND DON M. YOST
Gates and Crellin Laboratories of Chemistry, California Institute of Technology, Pasadena, California
November 16, 1940

In a recent review of the experimental infra-red and Raman spectra of ethylene and tetradeuteroethylene,[1] Conn and Sutherland have succeeded in giving very reasonable estimates of the positions of the four fundamental levels associated with the "rocking frequencies" of the methylene groups in C_2H_4. They have at the same time revised the previous estimate of the position of the "torsional" level[2] to 700 cm^{-1}, for reasons which are in our opinion inadequate, and with results which prove to be unsatisfactory. In particular, as they themselves remark, the calculated heat capacities then become too large. We wish to point out that a return to the Bonner assignment of this level not only yields a more reasonable explanation of the intensities of the weak combination and harmonic lines in the Raman spectrum, but also leads to a theoretical heat capacity which is in good agreement with the existing measured values[3] and with new experimental values recently obtained in these Laboratories.

In the Conn and Sutherland analysis, the value 700 cm^{-1} for the fundamental A_{1u} level[4] was a consequence of three vibrational assignments. The upper state of the Raman line at 1656 cm^{-1} was assigned as the binary combination state $(700)+(950)B_{3g}$, that of the Raman line at 3272.3 cm^{-1} as $2(700)+2(950)A_{1g}$, and that of the rather doubtful prism infra-red band of Coblentz at about 1720 cm^{-1} as $(700)+(1030)B_{3u}$. We believe that it is equally satisfactory to assign these levels in terms of an A_{1u} fundamental at about 829 cm^{-1} as $2(829)A_{1g}$, $(1623)+2(829)A_{1g}$, and $(829)+(943)B_{1u}$, respectively. That the frequency agreement for the Coblentz band is less close with this assignment seems of small importance, since neither its position nor its reality is certain. The intensity of the 1656-cm^{-1} Raman line may now be attributed to partial vibrational resonance of its upper state with the very strongly Raman active fundamental level $(1623.3)A_{1g}$, and the 3272.3 cm^{-1} line, while retaining its identity as a member of the 3272.3, 3240.3 cm^{-1} resonance doublet, is reduced from a quaternary to a ternary combination line.

TABLE I. *The heat capacity of ethylene gas.*

TEMPERATURE (°K)	$C_p°$ OBSERVED (CAL./MOLE-DEGREE)	MEAN DEVIATION (%)
270.7	9.74	0.1
300.0	10.39	0.4
320.7	10.99	0.2

The heat capacity of ethylene gas (Ohio Chemical Co., 99.5 percent C_2H_4) has been determined at three temperatures by the adiabatic expansion method of Lummer and Pringsheim, using pure nitrogen, whose heat capacity was calculated from spectroscopic data, as the standard comparison gas. In Table I are given average experimental values of C_p^0, the molal heat capacity at constant pressure corrected to infinite attenuation, together with the mean deviations of the individual determinations from the average. According to the experience of the authors in measurements on gases of known heat capacities these values may be expected to be accurate to within ±0.5 percent. It was assumed that the gas obeys the equation of state

$$P=RT/(V-B_0), \qquad (1)$$

where the coefficient B_0 is given by [5]

$$B_0=121.6-7.50\times10^4(1/T)-2.26\times10^8(1/T^3). \qquad (2)$$

The experimental values presented in Table I, together with the older data,[3] are shown graphically in Fig. 1. The dotted curve presented in this figure shows the heat capacity of ethylene calculated to the harmonic oscillator approximation with the Conn and Sutherland vibrational assignments, while the solid curve shows the similarly calculated heat capacity with the "torsional" level placed at 828.5 cm^{-1}. Deviations of the experimental points from the dotted curve are clearly too large and too systematic to be the result of experimental error, while the solid curve is in excellent agreement with the measured values.

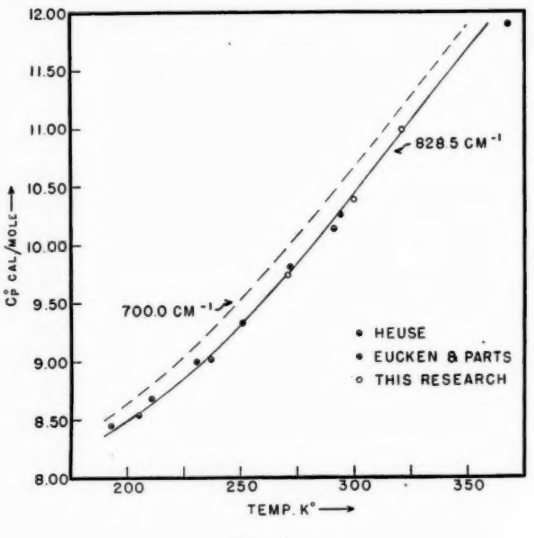


FIG. 1.

In the range $1F5$–$4F5$°K, these theoretical heat capacities in cal./mole-degree may be represented to within 0.02 by the following empirical equation:

$$C_p^\circ = 9.638 - 3.109\times10^{-3}T + 1.551\times10^{-4}T^2 - 1.426\times10^{-7}T^3. \quad (3)$$

Until the details of the 10μ band in ethylene are completely understood, the following fundamental levels, used in calculating the full curve of Fig. 1, may be accepted as essentially correct:

A_{1g};	3019.3	B_{1u};	3105.7	B_{2u};	2988.2	B_{3g};	3075.0
	1623.3		1027.5		1443.9		1021.5
	1342.4						
B_{3u};	948.8	B_{1g};	943.0	A_{1u};	828.5.		

[1] Conn and Sutherland, Proc. Roy. Soc. **A172**, 172 (1939).
[2] Bonner, J. Am. Chem. Soc. **58**, 34 (1936).
[3] Eucken and Parts, Zeits. f. physik. Chemie **B20**, 184 (1933); Heuse, Ann. d. Physik **59**, 86 (1919).
[4] These group-theoretical designations are uniquely defined by the standard tables of irreducible representations of the crystallographic point groups and the assertion that the z axis has been chosen to lie along the carbon-carbon bond, and the x axis to lie in the plane of the molecule.
[5] Keyes, J. Am. Chem. Soc. **60**, 1761 (1938).

Simultaneous Determination of Adiabatic and Isothermal Elasticities

H. F. LUDLOFF
School of Engineering, Cornell University, Ithaca, New York
December 1, 1940

Simultaneous measurement of the temperature curve of an isothermal and a corresponding adiabatic elastic modulus through the temperature region of phase transitions of higher order seems possible by any dynamical method using a proper frequency system, one part of which consists of isothermal, the other part of adiabatic proper vibrations. To determine appropriate conditions for such a method, the following problem has been investigated:

How does the adiabatic or isothermal character of a proper vibration depend on the frequency range, the vibrational type and the form of the sample?

From the investigations of Herzfeld and Rice[1] one can conclude that also in solids, plane sound waves above about 10^{11} cycles/sec. are to be considered as isothermal, while waves of lower frequency should be adiabatic.

However, consideration of proper vibrations of solids shows that somewhere in the low frequency range a change must occur, since for frequency 0, i.e., under static conditions, isothermal behavior is to be expected again. It is rather evident that the temperature distribution existing in a plane wave will be changed, when the boundary conditions begin to play a part. This is the case, when the distance between temperature maxima and minima becomes comparable with one of the dimensions of the sample. That again depends on the type of the excited vibration: For longitudinal vibrations this distance is characterized by the wave-length; for transverse vibrations successive regions of compression can only occur normally to the "direction of propagation." In the case of bending vibrations of plates, reeds, rods and circular wire loops the temperature gradient across the thickness is the only important one.

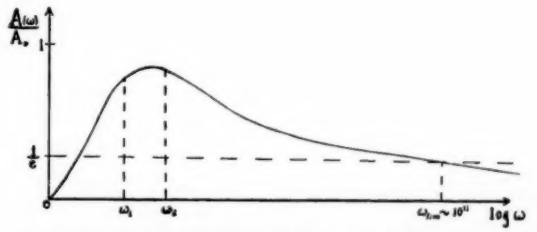


FIG. 1. Case A: all three dimensions of the sample of the same order. The $A(\omega)$ curve refers to longitudinal, transverse and surface vibrations, which in a finite solid occur combined.

The temperature distribution in a vibrating sample can be expressed by:

$$T = T_0 + (A_1(\omega) + iA_2(\omega)) \cdot f(xyz) \cdot e^{i\omega t}.$$

The function T is characterized by two temperature vibrations, one of which is in phase, the other out of phase with the elastic vibrations. If the absolute value of the temperature amplitude $|A_1+iA_2| = A(\omega)$ is small, the vibration is isothermal, otherwise it is adiabatic. $A(\omega)$ has to be normalized by the "perfectly adiabatic" amplitude A_0.

Taking into account the processes of heat conduction and heat production or absorption, and inserting the boundary conditions imposed by the size of the sample, we computed $A(\omega)$ for the different types of vibrations and found the following results:

(a) If the three dimensions of the sample are all of the same order, $A(\omega)$ turns out to be of the form given in Fig. 1. The $A(\omega)$ curve refers to longitudinal, transverse and surface vibrations, which in a finite solid occur combined. One sees that isothermal conditions (below the dotted line) exist above 10^{11} cycles/sec. and in the low frequency range only below the first proper frequency. There are no isothermal proper vibrations here, only forced isothermal vibrations.

(b) If one of the sample dimensions is small compared to the other two (plates, reeds, rods, loops), one has the situation given in Fig. 2. The $A(\omega)$ curve of Fig. 1 splits into two $A(\omega)$ curves characterizing vibrations, which are symmetrical or antisymmetrical with respect to the longitudinal mid-plane of the sample; finally the symmetrical vibrations are shown to pass over into compressional ones, the antisymmetrical ones pass over into bending vibrations. All of the symmetrical vibrations are adiabatic, as in the

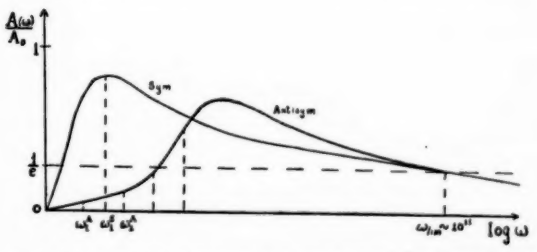


FIG. 2. Case B: one of the sample dimensions small compared to the other two. The $A(\omega)$ curves refer to symmetrical and to antisymmetrical vibrations.

case A. Of the antisymmetrical (bending) vibrations the first few are isothermal, the number depending on the thickness of the sample.

With the antisymmetrical vibrations it can be arranged that for appropriate dimensions of the sample the fundamental tone produces an isothermal curve and the first overtone a relatively "adiabatic" temperature curve of the elastic parameter.

These situations are being investigated experimentally for the order-disorder transformation of certain alloys, e.g. beta-brass.

Even for metals at room temperature the difference between the moduli might be determined by such a method.

[1] Herzfeld and Rice, Phys. Rev. 31, 691 (1928).

The Reaction of Hydrogen and Oxygen in the Presence of Silver. The Third Explosion Limit*

HAROLD R. HEIPLE** AND BERNARD LEWIS†
Bureau of Mines, U. S. Department of the Interior, Washington, D. C.
October 8, 1940

The experiments of Hinshelwood, Moelwyn-Hughes and Rolfe[1] with quiescent mixtures of hydrogen and oxygen in a silver vessel up to 700°C show that the chain reaction is virtually suppressed. Even at temperatures where the reaction rate has become very rapid or even explosive in a silica vessel, only a slow catalytic surface reaction is observed in the silver vessel. If chain initiation takes place in the gas phase that must be assumed at the above high temperature, these facts cannot be explained by a presumably very high chain-breaking efficiency of the silver surface.[2] Rather, chain carriers would have to be destroyed in the gas phase by the action of silver or some silver compound in the vapor form.[2]

To test this point directly a coil of silver ribbon or pieces of silver were mounted in the center of a tube, and the tube was heated to 625°-660°C, where the vapor pressure of silver is negligibly small. When hydrogen was passed through the tube for one hour no deposit of silver was visible on the glass tube. When oxygen was passed through for one hour an orange-brown coloration (blue-violet by reflected light) appeared on the tube; the deposit was distributed more or less evenly. With a mixture of hydrogen and oxygen for 20 to 30 minutes a heavy greenish-brown deposit having a metallic luster appeared on the tube in the form of a sharp outline of the spiral or silver pieces. When nitrogen was passed through for one hour, only a very slight deposit was formed. The deposited silver appeared to have penetrated the glass, for it could not be dissolved. Silver was determined spectroscopically. It might be mentioned that a small porcelain tube carrying the thermocouple wires, which was mounted inside the spiral, had accumulated a gray deposit that was identified as silver by dissolving in nitric acid and determining in the usual way. In none of the experiments was appreciable silver deposited downstream. The deposition of silver by oxygen has been known for many years and is thought to be caused by the formation of a volatile oxide of silver that decomposes at the glass wall.[3] The deposit with hydrogen and oxygen might, however, be caused by a sputtering of silver atoms in the chemical reaction at the surface as has been suggested earlier.[2] Rather high concentrations of silver in the gas phase that could exert a powerful chain-breaking effect must therefore have been present.

If the ability to destroy chain carriers is ascribed to the silver in the gas phase these experiments remove the difficulty that the silver-vessel experiments presented to the theory.

We have been making an extensive study of the third explosion limit of hydrogen and oxygen with respect to the effect of vessel dimensions, nature of surface, inert gases, and temperature. The results are in the main consistent with the theory[4] of the third limit based upon chain-breaking at the wall and chain-branching, thus supporting the formerly expressed view[4] that the third limit is of the branched-chain type, with superimposed thermal effects.

* Published with permission of the Director, Bureau of Mines, U. S. Department of the Interior.
** Bureau of Mines-University of Pittsburgh Research Fellow.
† Physical chemist, Physical Chemistry Section, Central Experiment Station, Bureau of Mines, Pittsburgh, Pa.
[1] C. N. Hinshelwood, E. A. Moelwyn-Hughes and A. C. Rolfe, Proc. Roy. Soc. London **A139**, 521 (1933).
[2] B. Lewis and G. von Elbe, *Combustion, Flames and Explosions of Gases* (Cambridge University Press, 1938), p. 68 *et. seq.* Cf. von Elbe and Lewis, J. Am. Chem. Soc. **59**, 656 (1937).
[3] K. F. Plattner, *Die Metallurgischen Röstprocesse* (Freiberg, 1856), p. 121. Cf. Mellor, Vol. 3, p. 343.
[4] Reference 2, p. 43 *et seq.*

Thermal Conductivity of Liquids

M. RAMA RAO
Department of Physics, University of Mysore, Bangalore, India
September 11, 1940

The thermal conductivity of liquids decreases with rise in temperature in contrast to the familiar and well-explained increase of gaseous thermal conductivity with temperature. The mechanisms of thermal conduction in liquids[1] so far proposed do not yield results in agreement with observed values. As a result the study has hitherto consisted largely of a collection of more or less well-established empirical equations of limited scope.

In a comparative and systematic study of the properties of the liquid state we[2] were led to define a reduced temperature $(\theta_c-\theta)/(\theta_c-\theta_s)$ where θ_c and θ_s refer to the critical and the freezing temperature and θ any temperature. Expressed in terms of this reduced temperature, the properties of the liquid could be written as $Z=Z_b\Theta^n$ where Θ is the reduced temperature defined above, Z is the property of the liquid at any temperature θ and Z_b the value at the melting point, and n any number. This reduced temperature has, however, failed to yield results of any value when dealing with a property such as viscosity. Applying the same considerations to the thermal conductivity of a liquid, it is found that neither the reduced temperature θ/θ_c nor the function Θ is useful in yielding any result. Assuming that a liquid is an assemblage of a number of linear harmonic oscillators each vibrating with a

frequency ν about a displaced equilibrium position and assuming that communication of thermal energy takes place when there is contact, between molecules, of very short duration, it is possible to arrive at an expression for the thermal conductivity of a liquid. When the temperature is increased the distribution of energy of molecules is governed by the Maxwell distribution law. If we make the simplifying assumption that communication of thermal energy between molecules now takes place when the difference of energy of the two attracting molecules has a constant magnitude, then it follows that the number of transfers of thermal energy will be decreased and hence the coefficient of thermal conduction at any temperature θ is given by $K=Ae^{B/\theta}$ where A and B are constants independent of temperature and θ is the absolute temperature of measurement. The relation is quite general and holds over the whole temperature range with a great variety of substances. One of the most interesting results (to be verified) which follows from this law of variation of K is that if two liquids be examined throughout a sufficient range of temperatures, a number of temperatures θ_1 can be found for liquid I at which the coefficient of conductivity is the same as it is for the other liquid II at a number of temperatures θ_2. If θ_1/θ_2 is plotted against θ_1 a straight line ought to result.

The significance of the constants A and B in relation to the properties of the liquid and other details will be published very soon.

[1] Bridgman, Proc. Am. Acad. **59**, 141 (1923); Kardos, Forsch Arb, Ingwes. **5**, 14 (1934); Smith, Trans. Am. Soc. Mech. Eng. **58**, 719 (1936).
[2] Sibaiya and Rama Rao, Cur. Sci. **8**, 8 (1939).

A Note on the Entropy of Fusion of Argon

O. K. Rice
University of North Carolina, Chapel Hill, North Carolina
November 27, 1940

Recently I have attempted to develop the idea that the disorder which arises when a crystal of argon melts is associated with a change of the average coordination number of the atoms from the value of 12 which holds for the crystal lattice.[1] A calculation of the associated entropy change was lacking, however, and it is the purpose of this note to indicate how this hiatus may be filled by a method based on the simplest possible considerations, and to present a preliminary test of this method.

In estimating the energy of disorder it was assumed that a coordination number less than 12 could be attained by removal of a part of the atoms from random lattice places, leaving holes, followed by a rearrangement without change in the density, but with division of the holes into smaller holes. The entropy change is closely connected with this rearrangement process.

Suppose, first, that there are N atoms and no holes. Then, by the Debye theory, the partition function per atom (for the thermal motion) is given by

$$Z_{12}=e(T/\Theta)^3 \qquad (1)$$

if the temperature is high enough to use classical mechanics, which is practically the case at the melting point. Θ is the characteristic temperature, a function of the interatomic distance. The subscript on Z denotes the coordination number.

Now suppose that there are N atoms and n holes, which gives a coordination number of $12N/(n+N)$. It is assumed that the atoms in any given row containing a hole are free to move into the hole and have their free volume increased by a corresponding amount. A hole can presumably be filled in along any particular direction, but once it is used for that direction it cannot simultaneously be used for another direction. We assume that one-third of the holes contribute to the free space in each of the three directions of space, and that the use of these holes in one direction is independent of the use of other holes in the other directions. The extra space made available in one direction will then be, roughly, $a/3$ per hole (where a is the interatomic distance) or $\frac{1}{3}(n/N)a$ per atom. If the atoms can be distributed in any way in this space, the contribution per atom to the partition function will be,[2] for each direction, $\frac{1}{3}(2\pi mkT/h^2)^{\frac{1}{2}}e(n/N)a$. (The factor e means that communal entropy is included for this part of the free volume. Actually this means that communal entropy contributes part of the entropy of fusion; this must be reserved for later discussion.) To get the total partition function for one direction we add $Z_{12}^{\frac{1}{3}}$, and cubing the result gives the complete partition function for the three directions:

$$Z_{12N/(n+N)}=[e^{\frac{1}{3}}T/\Theta+\tfrac{1}{3}(2\pi mkT/h^2)^{\frac{1}{2}}e(n/N)a]^3. \qquad (2)$$

This expression (which is, of course, for the thermal energy only) could be combined with a term arising from the potential energy, as previously obtained,[1] to give a partition function which would provide a complete theory of the fusion process. Whether this will fit the experimental data or not, and how it compares with other theories, must be determined by future calculations. In the meantime, however, it is possible to estimate n/N and ΔS, the entropy of fusion, if we take the experimental value for ΔV of fusion,[3] the x-ray data for a for the liquid,[4] and the density of the solid at the melting point,[3] and estimate[5] Θ as a function of a. We get $n/N=0.199$ for the liquid (which gives a coordination number of 10.0, agreeing with Miller and Lark-Horovitz[4]), and $\Delta S=3.5$ assuming no disorder in the solid. The experimental value is 3.35. Work now in process indicates that there is, perhaps, 0.6 entropy unit of disorder in the solid at the melting point. A correction for this brings ΔS down to 3.1. The results are rather sensitive to the values of a for liquid and solid. The agreement makes it seem worth while to work out the complete theory of fusion on this basis. Such work is in progress.

[1] O. K. Rice, J. Chem. Phys. **7**, 136, 883 (1939); see also J. D. Bernal, Trans. Faraday Soc. **33**, 27 (1937).
[2] See L. Tonks, Phys. Rev. **50**, 955 (1936).
[3] K. Clusius and K. Weigand, Zeits. f. physik. Chemie **B46**, 1 (1940).
[4] E. P. Miller and K. Lark-Horovitz, Phys. Rev. **58**, 207 (1940); Nature **146**, 459 (1940); but see A. Eisenstein and N. S. Gingrich, Phys. Rev. **58**, 307 (1940).
[5] O. K. Rice, Columbia Symposium paper, J. Am. Chem. Soc. in press.